

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 07:02:14 ; Search time 7880 Seconds

(without alignments)
10918.264 Million cell updates/sec

Title: US-09-890-813-5

Perfect score: 1985
Sequence: 1 gcaccagacagacagagag.....ataaacattatcatatct 1985

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBdb1: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pac: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vi: *
30: em_hcg_hum: *
31: em_hcg_in: *
32: em_hcg_other: *
33: em_hcg_mus: *
34: em_hcg_pln: *
35: em_hcg_rtd: *
36: em_hcg_mam: *
37: em_hcg_vrt: *
38: em_sy: *
39: em_hcgo_hum: *
40: em_hcgo_mus: *
41: em_hcgo_other: *

Result No. Score Query Match Length DB ID Description

1	1983	99.9	1985	6	AX180344	AX180344 Sequence
2	1863.2	93.9	1953	6	AX180346	AX180346 Sequence
3	1062.8	53.5	2037	8	AK073189	AK073189 Oryza sat
4	798	40.2	1991	8	AK061941	AK061941 Oryza sat
5	791.4	39.9	1343	8	AB042521	AB042521 Oryza sat
6	787.6	39.7	1916	8	AK121930	AK121930 Oryza sat
7	777.6	39.2	1983	8	AK102162	AK102162 Oryza sat
8	742.2	37.4	1695	8	AX135862	AX135862 Glycine m
9	707.6	35.6	1710	6	AX507049	AX507049 Sequence
10	707.6	35.6	1710	8	BT000493	BT000493 Arabidops
11	707.6	35.6	1971	8	AY057674	AY057674 Arabidops
12	706	35.6	1980	8	ATAXLYS1	X98873 A.thaliana
13	695.6	35.0	1933	8	AY089207	AY089207 Arabidops
14	678.4	34.2	1923	8	ATU62020	U62020 Arabidops
15	674.6	34.0	1658	6	AX180354	AX180354 Sequence
16	674.6	34.0	1658	6	BT009484	BT009484 Triticum
17	654.6	33.0	2164	8	AY088366	AY088366 Arabidops
18	555	28.0	565	6	AX180340	AX180340 Sequence
19	448.6	22.6	513	6	AX180342	AX180342 Sequence
20	371.2	18.7	847	6	AX180350	AX180350 Sequence
21	244.2	12.3	282	6	AR247449	AR247449 Sequence
22	225.2	11.3	455	6	AX180348	AX180348 Sequence
23	204.2	10.3	455	6	AX180352	AX180352 Sequence
24	158.2	7.0	160925	8	AP004332	AP004332 Oryza sat
25	139.4	7.0	167881	8	AP004330	AP004330 Oryza sat
26	129.8	6.5	10808	1	U67506	U67506 Methanococ
27	129.8	6.5	11000	6	AR271569_09	Continuation (10 O
28	129.8	5.8	9800	1	AE010670	AE010670 Methanosa
29	114.4	5.6	11629	1	AE010310	AE010310 Methanopy
30	111.8	5.5	148354	8	AP006529	AP006529 Oryza sat
31	108.4	5.4	13644	1	AE013376	AE013376 Methanosa
32	106.8	5.4	82612	8	AT731B5	AT163491 Arabidops
33	106.6	5.1	13494	1	AE001056	AE001056 Archaeogl
34	101.8	5.1	133846	8	AC092559	AC092559 Oryza sat
35	101	5.1	145290	1	AE015944	AE015944 Clostridi
36	92.2	4.6	30205	1	AE016935	AE016935 Bacteroid
37	91.4	4.6	302960	8	ATY16255	Y16255 Arabidops
38	90.4	4.6	4952	8	AB007650	AB007650 Arabidops
39	90.4	4.6	65465	8	AB007650	AB007650 Arabidops
40	89.8	4.5	301442	1	AE004127	AE004127 Vibrio ch
41	87.4	4.4	13866	1	AP005332	AP005332 Vibrio vu
42	86.6	4.3	13315	1	AE007541	AE007541 Clostridi
43	86.2	4.3	89154	8	ATAC010797	ATAC010797 Arabidops
44	86.2	4.3	89154	8	ATAC010797	Arabidops
45	86	4.3	89154	8	ATAC010797	Arabidops

ALIGNMENTS

RESULT 1
AX180344 1985 bp DNA linear PAT 06-AUG-2001
LOCUS AX180344
DEFINITION Sequence 5 from Patent WO0146393.
ACCESSION AX180344
VERSION AX180344.1 GI:15132291
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

Wed Mar 24 14:12:53 2004

us-09-890-813-5.rge

Page 2

```

JOURNAL      Patent: WO 0146393-A 5 28-JUN-2001;
              E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
  source
    1..1985
    /organism="Nea mayas"
    /mol_type="unasigned DNA"
    /db_xref="taxon:457"
ORIGIN

```

Query Match	99.9%	Score 1983	DB 6	Length 1985
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1985	Conservative	0	Mismatches	0
			Indels	Gaps
QY 1 GCACCGACACGAGCAGAGAGACTGCAAAATGCAATCCCAATGCGAGTCGCGTCCCGCGCC				60
Db 1 GCACCGACACGAGCAGAGAGACTGCAAAATGCAATCCCAATGCGAGTCGCGTCCCGCGCC				60
QY 61 CCGCGCGCTCGTTCCTCGATACCTCCCGCGAGCTCTGAGCATGTTGAGAGATTCGCGTGG				120
Db 61 CCGCGCGCTCGTTCCTCGATACCTCCCGCGAGCTCTGAGCATGTTGAGAGATTCGCGTGG				120
QY 121 CTTCCGTAACCCGAAACCGGGCTCGCGGTCGCAAGAGGGTGTCAATGCTGCTCGCCACATC				180
Db 121 CTTCCGTAACCCGAAACCGGGCTCGCGGTCGCAAGAGGGTGTCAATGCTGCTCGCCACATC				180
QY 181 CACACAGCCGTCCGGGCCCAACGAGCGAGCGCGGGACCGCGCTCTTGGGGCGCTGTCTT				240
Db 181 CACACAGCCGTCCGGGCCCAACGAGCGAGCGGGACCGCGCTCTTGGGGCGCTGTCTT				240
QY 241 CGAGAGGCTTCGGGATGAGAGGATTTGGGGGATACAGTCAGCGTGGTGAATGAGTTCCGGGG				300
Db 241 CGAGAGGCTTCGGGATGAGAGGATTTGGGGGATACAGTCAGCGTGGTGAATGAGTTCCGGGG				300
QY 301 GTCTCGGTCTCGTCGGCGCGAGAGATGGCTGAGTGGCGGCTCATCTTGACGTTCCC				360
Db 301 GTCTCGGTCTCGTCGGCGCGAGAGATGGCTGAGTGGCGGCTCATCTTGACGTTCCC				360
QY 361 CGAGAGAGCGCCGCTGCTGTCCTCTCTGAGCAATGGGGAAACACACCAACCAACCTCTGCT				420
Db 361 CGAGAGAGCGCCGCTGCTGTCCTCTCTGAGCAATGGGGAAACACACCAACCAACCTCTGCT				420
QY 421 TGCTGAGAGAAAGCAGTAGAGGTGTGAGATTATCCATGTTCTGAAATCGAAGATGAGAA				480
Db 421 TGCTGAGAGAAAGCAGTAGAGGTGTGAGATTATCCATGTTCTGAAATCGAAGATGAGAA				480
QY 481 TATGCTTAAAGCTTCATATCAAGACGCTGAGTGAATCTTGAGCTTCCAGAGATCTGTAA				540
Db 481 TATGCTTAAAGCTTCATATCAAGACGCTGAGTGAATCTTGAGCTTCCAGAGATCTGTAA				540
QY 541 TACAAAGCTTATGAACCTGAGCACTATATGAAAGATTCGCTATGATGAAGAGCTGAC				600
Db 541 TACAAAGCTTATGAACCTGAGCACTATATGAAAGATTCGCTATGATGAAGAGCTGAC				600
QY 601 GCTTAGAGACAGAGTACTACTTGTTCATTTTGGAGAAATGATGTCACACAGAGATTTTTC				660
Db 601 GCTTAGAGACAGAGTACTACTTGTTCATTTTGGAGAAATGATGTCACACAGAGATTTTTC				660
QY 661 TGCTTATTTGACAAAATTTGTGTCAAGGCAAGGAGATGACGATTTGATTTGATTT				720
Db 661 TGCTTATTTGACAAAATTTGTGTCAAGGCAAGGAGATGACGATTTGATTTGATTT				720
QY 721 CATTAACAACGATGAATTTGTGATCCGATATCTTAAAGCAACCTATCTGCTGTGTC				780
Db 721 CATTAACAACGATGAATTTGTGATCCGATATCTTAAAGCAACCTATCTGCTGTGTC				780
QY 781 GAAGAGACTTCATGGGACTGATACAGATCCAGGCAATCCTGTGTTATCTGCGTTCCT				840
Db 781 GAAGAGACTTCATGGGACTGATACAGATCCAGGCAATCCTGTGTTATCTGCGTTCCT				840
QY 841 TGGAGAGGCTCGAAATCTGTGCTGTAACTACTTTTAGCCGAGAGTGGTATGACTTGAC				900
Db 841 TGGAGAGGCTCGAAATCTGTGCTGTAACTACTTTTAGCCGAGAGTGGTATGACTTGAC				900

QY	901	TGCTACAAACCATGGTAAAGCTTGGGACTGAGAGAAATCATCGATATGGAAGAATGTTGA	960
Db	901	TGCTACAAACCATGGTAAAGCTTGGGACTGAGAGAAATCATCGATATGGAAGAATGTTGA	960
QY	961	TGATGATCTTTCCTTGATCATCAAAATATCAACCATGCGAAAGATGTTCCATCTTAC	1020
Db	961	TGATGATCTTTCCTTGATCATCAAAATATCAACCATGCGAAAGATGTTCCATCTTAC	1020
QY	1021	ATTGGAAGGCCACAGAACTTGCTTATTTTGGTGGTCAGGTTTGGCATCCAGATGAT	1080
Db	1021	ATTGGAAGGCCACAGAACTTGCTTATTTTGGTGGTCAGGTTTGGCATCCAGATGAT	1080
QY	1081	GAGACCTGGTAGAGAGGTGATATCCAGTTAGGGTTAAGAAATTCATCAACCCTTAAAG	1140
Db	1081	GAGACCTGGTAGAGAGGTGATATCCAGTTAGGGTTAAGAAATTCATCAACCCTTAAAG	1140
QY	1141	TCGAGCAACCTTATTTACCGAGCAAAAGACATGATATANGSTCTGGTGTGATCTAATAG	1200
Db	1141	TCGAGCAACCTTATTTACCGAGCAAAAGACATGATATANGSTCTGGTGTGATCTAATAG	1200
QY	1201	CATAGGCTCAAGTCAAAATGCACATATGTTGACATTTGAGCACTCGGATGCTTGCTCA	1260
Db	1201	CATAGGCTCAAGTCAAAATGCACATATGTTGACATTTGAGCACTCGGATGCTTGCTCA	1260
QY	1261	GTATGGTTTTCTGGCAAGGATATAGGATTTTGCCTATATGAAATCTATGTATATCTGT	1320
Db	1261	GTATGGTTTTCTGGCAAGGATATAGGATTTTGCCTATATGAAATCTATGTATATCTGT	1320
QY	1321	GGAATGTGTGCTACCATGTAAGTATGTTTCTGTGCTACATTCGATCAAAAGATCTG	1380
Db	1321	GGAATGTGTGCTACCATGTAAGTATGTTTCTGTGCTACATTCGATCAAAAGATCTG	1380
QY	1381	GAGTAGGGAATGATATACGACGGGCAAGTGAACCTTGACCATGTATGTTGAAGGCTTGAGA	1440
Db	1381	GAGTAGGGAATGATATACGACGGGCAAGTGAACCTTGACCATGTATGTTGAAGGCTTGAGA	1440
QY	1441	AATAGCAATTTGCTCTACTTACGACAGAGGCGAATATTTCACTTATCGAAATGTGGA	1500
Db	1441	AATAGCAATTTGCTCTACTTACGACAGAGGCGAATATTTCACTTATCGAAATGTGGA	1500
QY	1501	GCAATGCTCTCGATACTAGAAAAAGCGGACGTGTGCTGAGGAAAAAGTGGGCTTAATGT	1560
Db	1501	GCAATGCTCTCGATACTAGAAAAAGCGGACGTGTGCTGAGGAAAAAGTGGGCTTAATGT	1560
QY	1561	TCAGATGATCTCGCAAGAGCGTCMAAGGTTAACATGTGCTGATAGTCATGATAGCGA	1620
Db	1561	TCAGATGATCTCGCAAGAGCGTCMAAGGTTAACATGTGCTGATAGTCATGATAGCGA	1620
QY	1621	TGCAAGGCACTGCTAGAAAGCCCTTCATCAGCGCTTCTTGAAAGAGATGCTTATCA	1680
Db	1621	TGCAAGGCACTGCTAGAAAGCCCTTCATCAGCGCTTCTTGAAAGAGATGCTTATCA	1680
QY	1681	AGTCGAAGCGAGAAACCTACTCGTGGGCTGATCAACGTAGGCTTTCTGGGTCAGAGGCT	1740
Db	1681	AGTCGAAGCGAGAAACCTACTCGTGGGCTGATCAACGTAGGCTTTCTGGGTCAGAGGCT	1740
QY	1741	GTTATCTGTTATAGATTCCTCACTGCTCCATGAACGGCATTGGGCAATGGATCATTTGATC	1800
Db	1741	GTTATCTGTTATAGATTCCTCACTGCTCCATGAACGGCAATGGGCAATGGATCATTTGATC	1800
QY	1801	ATGTTTGGCTGAAACAAGATATGTCTTCAGGTCTCAGGCAATGACTGCAAAACCTGTGT	1860
Db	1801	ATGTTTGGCTGAAACAAGATATGTCTTCAGGTCTCAGGCAATGACTGCAAAACCTGTGT	1860
QY	1861	TTCTGTTTGAACCTGTTTGACAGACCAAGTGAAGCTCGAGCACCGATTTGTCACAAAT	1920
Db	1861	TTCTGTTTGAACCTGTTTGACAGACCAAGTGAAGCTCGAGCACCGATTTGTCACAAAT	1920
QY	1921	GGCAAGCCTGTGATATATTCACATGTGCTCTAATCAATATATATATATTAACATATCA	1980
Db	1921	GGCAAGCCTGTGATATATTCACATGTGCTCTAATCAATATATATATATTAACATATCA	1980

Db 1981 TATCT 1985

RESULT 2
AX180346
LOCUS AX180346 1953 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 7 from Patent WO0146393.
ACCESSION AX180346
VERSION AX180346.1 GI:15132292
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoidae; Andropogoneae; Zea.

REFERENCE
1 Falco, S.C., Famodu, O.O. and Thorpe, C.J.
Aspartate Kinase
Patent: WO 0146393-A 7 28-JUN-2001;
JOURNAL E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
source location/Qualifiers
1.1953
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match	93.9%	Score 1863.2	DB 6	Length 1953
Best Local Similarity	98.4%	Pred. No. 0		
Matches 1941	Conservative 0	Mismatches 8	Indels 23	Gaps 5

QY 11 GAGCAGAGAGACTCGAAGATGCAATCCAGTGGATCGCTCGCGCCGCCGCCCTC
DB 1 GAGCAGAGAGACTCGAAGATGCAATCCAGTGGATCGCTCGCGCCGCCGCCCTC 60
QY 71 GTTCGCTGATACCTCCGCGGAGCTCGACATGTTGAGAGACTGGCGCTGCTCGGTC
DB 61 GTTCGCTGATACCTCCGCGGAGCTCGACATGTTGAGAGACTGGCGCTGCTCGGTC 120
QY 131 CGAACCGGCGCTCGCGATGCAAGAGGTTGCAATGGTGGTGGCGGACTCCACGAGCT
DB 121 CGAACCGGCGCTCGCGATGCAAGAGGTTGCAATGGTGGTGGCGGACTCCACGAGCT 180
QY 191 CGGCGCAAGCAAGCGGAGCGGCGGAGCGGCTGCTGGGCGGCTGTTCTCGAGGCGCTC
DB 181 CGGCGCAAGCAAGCGGAGCGGCGGAGCGGCTGCTGGGCGGCTGTTCTCGAGGCGCTC 240
QY 251 GGGATGAGAGGATGAGGAGATCACTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
DB 241 GGGATGAGAGGATGAGGAGATCACTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 300
QY 311 TCGTGGCGCGGAGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
DB 301 TCGTGGCGCGGAGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 360
QY 371 CCCGTCGCTGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
DB 361 CCCGTCGCTGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 431 AAGGCAAGAGGAGTGGAGTATCATGTTTCTGAAATCGAAGAGTGAATGTCGAAA
DB 421 AAGGCAAGAGGAGTGGAGTATCATGTTTCTGAAATCGAAGAGTGAATGTCGAAA 480
QY 491 AGCTTACATTCAGAGCGGTGAGTGAATCTGGAATCTGGAATCTGGAATCTGGAATCTG
DB 481 AGCTTACATTCAGAGCGGTGAGTGAATCTGGAATCTGGAATCTGGAATCTGGAATCTG 539
QY 551 T---ATGAACTGAGCACTATTGAAGGATGCTATGATGAAGAGCTACGCTAG 606
DB 540 GCTAGATGATGCTGAGCAACTATTGAAGGATGCTATGATGAAGAGCTACGCTAG 599
QY 607 GACGAGTACTACTGTTTCTATTGGAGAAATGATGTCACCAAGATTTTCTGCTTA 666

Db 600 GACGAGTACTACTGTTTCTATTGGAGAAATGATGTCACCAAGATTTTCTGCTTA 659
QY 667 TTGAAACAAAATTTGGTGTCAAGGACGGGAGTATGAGCACTTGAATGTTGTTCAATC 726
DB 660 TTGAAACAAAATTTGGTGTCAAGGACGGGAGTATGAGCACTTGAATGTTGTTCAATC 719
QY 727 AACTGATGATTTGGTAAATGCGATATCTTGAAGCAACTATCTGCTGTTGCGAAG 786
DB 720 AACTGATGATTTGGTAAATGCGATATCTTGAAGCAACTATCTGCTGTTGCGAAG 779
QY 787 ACTTATGAGGAGTGGATGATACAGATCCAGGATACCTGTTACTGAGGTTCTTGAAG 846
DB 780 ACTTATGAGGAGTGGATGATACAGATCCAGGATACCTGTTACTGAGGTTCTTGAAG 839
QY 847 GGGCTGAAATCTGGTCTGTTACTACTTTAGCCGAGGTGATGACTGACTGCTAC 906
DB 840 GGGCTGAAATCTGGTCTGTTACTACTTTAGCCGAGGTGATGACTGACTGCTAC 899
QY 907 AACCAATGAGTAAAGCTTGGAGTGAAGAAATCAAGTATGGAAGATGTTATGTTGT 966
DB 900 AACCAATGAGTAAAGCTTGGAGTGAAGAAATCAAGTATGGAAGATGTTATGTTGT 959
QY 967 ACTTACTGTGATCCAAATATCTAACCAATGCAAGACCTGTTCACTTAAATTTGA 1026
DB 960 ACTTACTGTGATCCAAATATCTAACCAATGCAAGACCTGTTCACTTAAATTTGA 1019
QY 1027 AAGGCGACAGAACTGTTGTTATTTGGTGGTCAAGTTTGCATCCAAATGATGAGAC 1086
DB 1020 AAGGCGACAGAACTGTTGTTATTTGGTGGTCAAGTTTGCATCCAAATGATGAGAC 1079
QY 1087 TGTGAGAGAGTGTATATTCAGTTAGGTTAGAAATTCATACACCTTAAAGCTTCAG 1146
DB 1080 TGTGAGAGAGTGTATATTCAGTTAGGTTAGAAATTCATACACCTTAAAGCTTCAG 1139
QY 1147 CACCTTATATACAGACAAAGACATGATGAAGTGGTGGTGTACTTAACTAGATAGT 1206
DB 1140 CACCTTATATACAGACAAAGACATGATGAAGTGGTGGTGTACTTAACTAGATAGT 1193
QY 1207 GCTCAAGTCAAAATGCTACTATGTTGGAATGTTGAGCATGCGATGCTGTCAGTATG 1266
DB 1194 GCTCAAGTCAAAATGCTACTATGTTGGAATGTTGAGCATGCGATGCTGTCAGTATG 1253
QY 1267 TTTTCTGGAGAGGATCAGATTTGCTATATGAAAGATCTATGATATCTGATGATG 1326
DB 1254 TTTTCTGGAGAGGATCAGATTTGCTATATGAAAGATCTATGATATCTGATGATG 1307
QY 1327 TGTGCTACAGTGAATGATGTTTCTGTTGCTGATGATGATGATGATGATGATG 1386
DB 1308 TGTGCTACAGTGAATGATGTTTCTGTTGCTGATGATGATGATGATGATGATG 1367
QY 1387 GGAATGATACAGCAGGCAATGAACTGACATGATGATGATGATGATGATGATG 1446
DB 1368 GGAATGATACAGCAGGCAATGAACTGACATGATGATGATGATGATGATGATG 1421
QY 1447 AATTGTTGCTACTTACAGCAGGCGATATTTCACTTATGGAATGATGATGATG 1506
DB 1422 AATTGTTGCTACTTACAGCAGGCGATATTTCACTTATGGAATGATGATGATG 1481
QY 1507 GTCTGTGATACAGGAGAGCGGAGCTGCTGAGAGAGAGAGAGAGAGAGAGAG 1566
DB 1482 GTCTGTGATACAGGAGAGCGGAGCTGCTGAGAGAGAGAGAGAGAGAGAGAG 1541
QY 1567 GATCTGCAAGAGAGGTTAAAGGTTAAATGATGATGATGATGATGATGATGATG 1626
DB 1542 GATCTGCAAGAGAGGTTAAAGGTTAAATGATGATGATGATGATGATGATGATG 1601
QY 1627 GGCATCTGTAAGAGCCCTTCACTAGGCGTTCTTTGAAGAGATGCTTATCAAGTGA 1686
DB 1602 GGCATCTGTAAGAGCCCTTCACTAGGCGTTCTTTGAAGAGATGCTTATCAAGTGA 1661
QY 1687 AAGCGAGAACTACTGCGGCTGATCAAGTATGCTGCTGCTGCTGCTGCTGCTG 1746

Db	1662	AGCGGAGAACTTACTGCTGAGGCTGATCAACGTAAGGCTTGTCTGAGCTCCAGGCGTGTATC	1721
Qy	1747	TGTTATAGATTCCTCCACTGCTGCATGAACGGGATGGGCGATTTGATCATGATGTT	1806
Db	1722	TGTTATAGATTCCTCCACTGCTGCATGAACGGGATGGGCGATTTGATCATGATGTT	1781
Qy	1807	TGCTTGAACCAAGTATGCTTCTGAGGCTTCAGCCCAATGACTGCAAACTGTGTTCTGT	1866
Db	1782	TGCTTGAACCAAGTATGCTTCTGAGGCTTCAGCCCAATGACTGCAAACTGTGTTCTGT	1841
Qy	1867	TTTAGAAGCTTTGGACAGACAGAGAGCTGGAGACCGATGTCACCAAGATGGCGAG	1926
Db	1842	TTTAGAAGCTTTGGACAGACAGAGAGCTGGAGACCGATGTCACCAAGATGGCGAG	1901
Qy	1927	CCTGTGATATTAATTCCTCACTGCTCTTAATCAATATATATTAATTAACATTATC	1978
Db	1902	CCTGTGATATTAATTCCTCACTGCTCTTAATCAATATATATTAATTAACATTATC	1953
RESULT 3			
AK073189		2037 bp	mRNA
LOCUS			linear
DEFINITION			FLN 24-JUL-2003
ACCESSION	AK073189		
KEYWORDS	FLN-CDNA, CAP		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE			
AUTHORS	1 The Rice Full-length cDNA Consortium, National Institute of		
	Agronomical Sciences Rice Full-length cDNA Project Team,		
	Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,		
	Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,		
	Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,		
	Onitsuki, K., Shishiki, T., Foundation of Advancement of International		
	Science Genome Sequencing & Analysis Group; Ootomo, Y., Murakami, K.,		
	Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,		
	Kuroseki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,		
	Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,		
	Ikedo, R., Ishibiki, J., Kawamata, M., Matsubara, K., RIKEN,		
	Kusunegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,		
	Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,		
	Hara, A., Hashidume, M., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,		
	Kagawa, I., Kondo, S., Komno, H., Miyazaki, A., Ootomo, N., Ota, Y.,		
	Saito, R., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,		
	Yoshino, M., and Hayashizaki, Y.		
	Collection, mapping, and annotation of over 28,000 cDNA clones from		
	Japonica rice		
TITLE	Science 301 (5631), 376-379 (2003)		
JOURNAL			
MEDLINE	22752273		
PubMed	12869764		
REFERENCE			
AUTHORS	2 (pages 1 to 2037)		
	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,		
	Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, M.,		
	Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,		
	Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,		
	Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,		
	Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,		
	Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,		
	Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Komno, H., Kouda, M.,		
	Koya, S., Kurihara, C., Kuroseki, T., Kusunegi, T., Li, C., Lu, M.,		
	Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,		
	Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,		
	Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,		
	Numaaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,		
	Ootomo, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,		
	Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,		
	Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,		
	Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,		

Qy	1662	AGCGGAGAACTTACTGCTGAGGCTGATCAACGTAAGGCTTGTCTGAGCTCCAGGCGTGTATC	1721
Db	1747	TGTTATAGATTCCTCCACTGCTGCATGAACGGGATGGGCGATTTGATCATGATGTT	1806
Qy	1807	TGCTTGAACCAAGTATGCTTCTGAGGCTTCAGCCCAATGACTGCAAACTGTGTTCTGT	1866
Db	1782	TGCTTGAACCAAGTATGCTTCTGAGGCTTCAGCCCAATGACTGCAAACTGTGTTCTGT	1841
Qy	1867	TTTAGAAGCTTTGGACAGACAGAGAGCTGGAGACCGATGTCACCAAGATGGCGAG	1926
Db	1842	TTTAGAAGCTTTGGACAGACAGAGAGCTGGAGACCGATGTCACCAAGATGGCGAG	1901
Qy	1927	CCTGTGATATTAATTCCTCACTGCTCTTAATCAATATATATTAATTAACATTATC	1978
Db	1902	CCTGTGATATTAATTCCTCACTGCTCTTAATCAATATATATTAATTAACATTATC	1953
TITLE			
JOURNAL			
COMMENT			
	URL: http://cdna01.dna.affrc.go.jp/cDNA/		
	NIBS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,		
	Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,		
	Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,		
	Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and		
	Yamamoto, M.		
	FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y.,		
	Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,		
	Kodama, T., Kuroseki, T., Kusunegi, T., Lu, M., Masuda, H., Miura, J.,		
	Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,		
	Sugiyama, A., Matsubara, K., and Murakami, K.		
	Yoshimura, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,		
	Genome Exploration Research Group in Riken: Adachi, J., Aizawa, K.,		
	and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,		
	Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,		
	Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K.,		
	Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,		
	Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,		
	Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Komno, H., Kouda, M.,		
	Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,		
	Nakamura, M., Nishi, K., Nomura, K., Numaaki, R., Ootomo, N.,		
	Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,		
	Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,		
	Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,		
	Takaku-Akahira, S., Tanaka, I., Tomaru, A., Toya, T., Waki, K.,		
	Yasunishi, A., and Hayashizaki, Y.		
FEATURES			
source	1..2037		
	/organism="Oryza sativa (japonica cultivar-group)"		
	/mol_type="mRNA"		
	/cuiFier="Nipponbase"		
	/db_xref="taxon:39947"		
	/clone="J033004110"		
ORIGIN			
Query Match	53.5%	Score 1062.8;	DB 8; Length 2037;
Best Local Similarity	80.3%	Pred. No. 1.8e-247;	
Matches 1314;	Conservative 0;	Mismatches 299;	Indels 23; Gaps 5;
Qy	96	CTGACATGTTGAGAGCTGCGCGTGTCTTGGATCCGACACCGGCGCTCGGATGACAAAG	155
Db	168	CAGCTCGTATCGTGTCTTGGAGAGGCTCCCTCCGCGCGCGCGAGGCGCGGA	227
Qy	156	GGTTCGATAGTGTGTCGCGACCTCCACAGCTCGGCGCGCAACGACGCGCGG	215
Db	228	GACTGAGGTGCGG	267
Qy	216	ACGGCGTCTTGGCG	275
Db	288	GGCG	347
Qy	276	TGAGGTGTGATGAAGTTCGGGCGGCTCGCGGTGTGTCGCGCGCGAGGATGCGAG	335
Db	348	TGAGGTGTGATGAAGTTCGGGCGGCTCGCGGTGTGTCGCGCGCGAGGATGCGAG	407
Qy	336	TGCGCGCGCTTACCTGACGTTCCCGAGAGCGCGCGCGCGCGCGCGCGCGCGCG	395
Db	408	TGCGCGCGCTTACCTGACGTTCCCGAGAGCGCGCGCGCGCGCGCGCGCGCGCG	467
Qy	396	GGAAACACCAACCACTTCTCTGCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAG	455

D	b	468	GGAAGACCAACAATCTCTCTCTCTCGCTGGGAGAAAGGACAGTGGATGGCGGGTATCC	527
Q	y	456	ATGTTTCTGAAATCCGAAGATGGAATATGCTCAAAAGCTACATATTAACACGGTGGATG	515
D	b	528	GTTGTTGGAAATTGAAGATGGAATTTGATCBAAGATCTACACATTTAAACCTGTGGAG	587
Q	y	516	AACCTGACTTCCAAAGMATCTGTAATACA---AGCCTTAATGAACTGAGCAACTATTG	571
D	b	588	AACCTGCACTGCCAAG-ATCTGTAAATACACACTATGCTGATGAACTGAGAGCAGCTCTG	646
Q	y	572	AAAGGATGCGATATGATGTAAGAAGTCAGCGCTAGGACCAAGTACATCCCTGTTTCAATT	631
D	b	647	AAAGGATTTGCANTGATGAAGAAAGCTGACCTTGGACACTGACCTACCTTGTTCATTT	706
Q	y	632	GGAAGATGATGTCCACAGATTTTTTCTGCTTAATTTGAACAATAATTCGTGTCAAGCA	691
D	b	707	GGGAGATGATGTCTACACGATTTTTTCTGCTTAATTTGAACAAATTCGAGTAAGAACA	766
Q	y	692	CGGCAATATGAGCGCAATTTGATATATGAGTTCAATTACAATGAAATTTGGTAAATGGGAT	751
D	b	767	CGGCAATATGAGCGCAATTTGATATATGAGTTCAATTACAATGAAATTTGGTAAATGGGAT	826
Q	y	752	ATCTTAAGAAACAACCTATCTGCTGTTGCGAAGAGACTTCAATGGGACATGGAATACAGAT	811
D	b	827	ATCTTAAGAAACAACCTATCTGCTGTTGCGAAGAGACTTCAATGGAAGACTGATTCGTGAT	886
Q	y	812	CCAGCATATACCTGTTGTTATCTGAGTTTCCTTGGAAAGGGCTGGAATTCGTGCTGTAAAT	871
D	b	887	CCAGCATATACCTGTTGTTATCTGAGTTTCCTTGGAAAGGGCTGGAATTCGTGCTGTAAAT	946
Q	y	872	ACTTTAGCCGAGGTGTGATGATCTGACCTGCTCAACAACATTTGGTAAAGCCTTGGGACTG	931
D	b	947	ACTTTAGCCGAGGTGTGATGATCTGACCTGCTCAACAACATTTGGTAAAGCCTTGGGACTG	1006
Q	y	932	AGGAAATTCAGGTAATGGAAGAATGTATGATGTACTCTTGTATCCAAATATCTAC	991
D	b	1007	AGGAAATTCAGGTAATGGAAGAATGTATGATGTACTCTTGTATCCAAATATCTAC	1066
Q	y	992	CCACATGCAAAAGCTGTTCATTAATTAAACATTTGAAGAGCCACAGAACTTGCTTATTTT	1051
D	b	1067	CCCAATGCAAAAGCTGTTCATTAATTAAACATTTGAAGAGCCACAGAACTTGCTTATTTT	1126
Q	y	1052	GGGCTCAGATTTTGATGCACAATTCAGATGAGACCTGTCGAGAAGGTATATTCAGAT	1111
D	b	1127	GGGCTCAGATTTTGATGCACAATTCAGATGAGACCTGTCGAGAAGGTATATTCAGAT	1186
Q	y	1112	AGGGTTAAGATTCATACACCTTAAGCTCCAGGACCCCTTATTAACAGACAAAGAGAC	1171
D	b	1187	AGGGTTAAGATTCATACACCTTAAGCTCCAGGACCCCTTATTAACAGACAAAGAGAG	1246
Q	y	1172	ATGATTAAGAGTCTGGTTGTACTACCTAGCTAGCTCAAGTCAAGTATGCTATATGTTG	1233
D	b	1247	ATGATTAAGAGTCTGGTTGTACTACCTAGCTAGCTCAAGTCAAGTATGCTATATGTTG	1307
Q	y	1232	GACATTTGAGCACTCGATGCTTGGTGTAGATATGTTTCTGGCAAGGGTATCAGGTAT	1291
D	b	1301	GATATATGTAAGCACTCGATGCTTGGTGTAGATATGTTTCTGGCAAGGGTATCAGGTAT	1366
Q	y	1292	TGCTATATTAAGAGTATATATCTGTGATGTTGTCTACACAGTGAAGTATGTTG	1351
D	b	1361	T-----TGAAGATCTAGGAGTATCTGTAAATTTGTGTGCTACTAGTGAAGTATG	1414
Q	y	1352	TCTGTGTCACTGTATTCATCAAAAGTCTGAGATGAGGAACTGATACAGCGCAATGAA	1411
D	b	1415	TCTGTGTGTGTATTCATCAAAAGTCTGAGAGGAGGAACTTATTAACAAG-----GAA	1466
Q	y	1412	CTTGACCAATGATGTAAGAGCTTGAAGAAATTAACAATCTTGTCTACTTCAGCAGAGG	1477
D	b	1469	CTTGACCAATGATGTAAGAGCTTGAAGAAATTAACAATCTTGTCTACTTCAGCAGAGG	1522
Q	y	1472	GCGATATTTCACTTATCGGAAATGTGAGCAATCGTCTGTGATCTAGAAAAAGCGGGA	1531
D	b	1529	GCAATATATCACTCTATGGAAGAGTACGGGATCATCTTATATCTGAAAAAGGCTTT	1586

Oy	153	CGGTGCTTAGGAAAGAGGGGGTAAAGTTCAAGTATCTCGCAAGAGACGTCGAAGTT	1591
Dd	1589	CAAGTCTTAGGAAAGAGGGGGTAAAGTTCAAGTATCTCGCAAGTGCATCCAGGTT	1648
Oy	1592	AACATCTGCCTGATAGTCCATGATAGCGATGCAAGGACACTGTAGAACCTTTATGAG	1651
Dd	1649	AACATCTCTCTATATGTTTCATGATAGGAGGCAAAACAGTGAATAAAGCCCTCCACAG	1708
Oy	1652	GGCTTTTGAAGACGATGTCCTTCAACAGTGGAGCGAGACCTACTCGTGGCGTGA	1711
Dd	1709	GGCTTTTGAAGAGAGATGTCCTTCAACAGTGGAGGAGAAAGCCCTTCACTAGCTGA	1768
Oy	1712	TCAACGTAGGCTTTGC	1727
Dd	1769	TCAGCTAGTATCCGCTAC	1784
RESULT 4			
AK061941		1991 bp	mRNA
LOCUS			linear
DEFINITION			Oriza sativa (Japonica cultivar-group) cDNA clone:001-042-D10, full insert sequence.
ACCESSION			AK061941
VERSION			AK061941.1
KEYWORDS			FLI; CDNA; oligo-capping.
SOURCE			Oriza sativa (Japonica cultivar-group)
ORGANISM			Oriza sativa (Japonica cultivar-group)
REFERENCE			Eukaryotic/Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oriza.
AUTHORS			1 The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team; Kishino,S., Satoh,K., Nagata,T., Kawasumi,N., Doi,K., Kishimoto,S., Yaaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Naito,T., Ohneda,E., Yahagi,W., Suzuki,K., Ii,C., Ohtsuki,K., Shisiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Oono,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurotaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Nakawara,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Mikiura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Mura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., SIKEN; Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Kono,H., Miyazaki,A., Oono,N., Ota,Y., Saito,R., Saeki,D., Sato,K., Shibata,K., Shingawa,A., Shiraki,T., Yoshino,M., and Hayashizaki,Y.
TITLE			Collection, mapping, and annotation of over 28,000 cDNA clones from Japonica rice
JOURNAL			Science 301 (5631), 376-379 (2003)
MEDLINE			22752273
PUBMED			12869764
REFERENCE			2 (bases 1 to 1991)
AUTHORS			Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanaoka,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoaka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kangawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawana,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kojima,T., Kojima,K., Kojima,Y., Kondo,S., Kono,H., Kouda,M., Koda,S., Kurihara,C., Kurotaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Mura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Naito,K., Naito,T., Naito,J., Naito,J., Nishikawa,K., Nomura,K., Numasaki,R., Ohneda,E., Oono,M., Ohtsuki,K., Oka,M., Ooka,H., Oono,N., Ota,Y., Oono,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Saeki,D., Saito,K., Satoh,K., Shibata,K., Shingawa,A., Shiraki,T., Shisiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,K., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahita,S., Tanaka,T., Tomaru,A.,

TITLE
JOURNAL

Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skkichi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28k full-length cDNA clones from japonica rice.

COMMENT

URL: http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Sato, K., Nagata, T., Kawasashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

PAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kuroaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Nishikura, Y., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory Group in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa, H., Kishikawa, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyama, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ono, M., Oseko, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku, Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES

Location/Qualifiers
1..1991
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultiVar="Nidpobare"
/db_xref="taxon:39947"
/clone="001-042-D10"

ORIGIN

Query Match

Best Local Similarity 74.2% Pred. No. 5e-183;
Matches 1053; Conservative 0; Mismatches 351; Indels 15; Gaps 3;

273 AGCTCAGCGTGTGATGAAAGTTGGGGGGCTCTGGTGTCTGTGGCGCGGAGATGCGT 332
Db AGCTCAGCGTGTGATGAAAGTTGGGGGGCTCTGGTGTCTGTGGCGCGGAGATGCGT 332
315 AGCTCAGCGTGTGATGAAAGTTGGGGGGCTCTGGTGTCTGTGGCGCGGAGATGCGT 374
QY AGCTGAGCGGCTCTCATCTGAGAGTCCCGAGAGAGCGCGCGTGTCTGTCTGTCTGCA 392
Db AGCTGAGCGGCTCTCATCTGAGAGTCCCGAGAGAGCGCGCGTGTCTGTCTGTCTGCA 392
375 AGTGGCCGATCTCTCTCACTTCCCGAGAGAGCGCGCGTGTCTGTCTGTCTGCA 414
QY TGGGGAAAACCAACCAACCTTCTCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 452
Db TGGGGAAAACCAACCAACCTTCTCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 452
435 TGGGGAGAGCAACCAACCAAGCTTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 494
QY TCCATGTTTCTGAATTCGAAGAGAGAGAGATGATGATGATGATGATGATGATGATG 512
Db TCCATGTTTCTGAATTCGAAGAGAGAGAGATGATGATGATGATGATGATGATGATG 512
495 CAAATGTGTGAG 554
QY ATGAACCTTGAGCTTCAAGATCTGTAAATACAGAGCTT---ATGAACCTGAGAGAG 569
Db ATGAACCTTGAGCTTCAAGATCTGTAAATACAGAGCTT---ATGAACCTGAGAGAG 569
555 ATCAACTTGAGATGAG 614
QY TGAAGAGTATGAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 629
Db TGAAGAGTATGAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 629
615 TTAAAGGAAATTCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674

QY TTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 689
Db TTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 689
675 TGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 734
QY CACGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 749
Db CACGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 749
735 CACGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 794
QY ATATCTTGAAGCAACCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 809
Db ATATCTTGAAGCAACCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 809
795 ATATCTTGAAGCAACCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 854
QY ATCAAGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 869
Db ATCAAGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 869
810 ATCAAGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 914
QY GTCTGCTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 914
Db GTCTGCTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 914
870 CTACTTGAAGCGAG 929
QY CCACTTGAAGCGAG 974
Db CCACTTGAAGCGAG 974
915 CCACTTGAAGCGAG 989
QY TGAAGAGATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 989
Db TGAAGAGATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 989
930 TGAAGAGATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1034
QY TGAAGAGATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1034
Db TGAAGAGATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1034
990 ACCAGATGCAAG 1049
QY ATCAAGTGAAGCAACCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1094
Db ATCAAGTGAAGCAACCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1094
1035 ATCAAGTGAAGCAACCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1154
QY TTGGTCTCAG 1154
Db TTGGTCTCAG 1154
1050 TTGGTCTCAG 1169
QY TTGGTCTCAG 1169
Db TTGGTCTCAG 1169
1095 TTGGTCTCAG 1214
QY TTGGTCTCAG 1214
Db TTGGTCTCAG 1214
1110 TTGGTCTCAG 1229
QY TTGGTCTCAG 1229
Db TTGGTCTCAG 1229
1170 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1268
QY ATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1268
Db ATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1268
1215 ATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1289
QY TGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1289
Db TGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1289
1230 TGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1322
QY TGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1322
Db TGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1322
1233 TGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1349
QY TGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1349
Db TGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1349
1233 TGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1409
QY TGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1409
Db TGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1409
1333 TATCTTGAAG 1442
QY TATCTTGAAG 1442
Db TATCTTGAAG 1442
1410 AACTTGAAG 1469
QY AACTTGAAG 1469
Db AACTTGAAG 1469
1443 AACTTGAAG 1502
QY AACTTGAAG 1502
Db AACTTGAAG 1502
1470 GGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1529
QY GGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1529
Db GGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1529
1503 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1562
QY GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1562
Db GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1562
1530 GACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1589
QY GACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1589
Db GACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1589
1553 TCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1622
QY TCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1622
Db TCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1622
1550 TTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1649
QY TTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1649
Db TTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1649
1623 TGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1682
QY TGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1682
Db TGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1682
1650 AGGCGTCTTGAAG 1688
QY AGGCGTCTTGAAG 1688
Db AGGCGTCTTGAAG 1688
1683 CAGCATCTTGAAG 1721

RESULT 5
AB042521 1343 bp mRNA linear PLN 13-MAY-2000
LOCUS
DEFINITION Oryza sativa mRNA for aspartate kinase, partial cds.
ACCESSION AB042521
VERSION AB042521.1 GI:7798568
KEYWORDS
SOURCE
ORGANISM
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta, Liliopsida; Poales; Poaceae;
Eriarctioideae; Oryzaceae; Oryza.
REFERENCE
1 (bases)
Kiyota, S.
Lysine sensitive aspartate kinase from rice
Published Only in Database (2000)
2 (bases 1 to 1343)
Kiyota, S.
Direct Submission
Submitted (09-MAY-2000) Seichiro Kiyota, National Institute of
Agrobiological Resources, Laboratory of Stress Physiology,
Kamondai 2-1-2, Tukuba, Ibaraki 305-8602, Japan
(E-mail:skiyota@abr.affrc.go.jp, Tel:+81-298-38-8382)
Location/Qualifiers
1..1343
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
<1..1065
/codon_start=1
/product="aspartate kinase"
/protein_id="BA95630.1"
/db_xref="GI:7798569"
/translation="EGECSTRTIPAIYKIKGVKARQYDADIFITTTDFGNADILE
AAYPAVKRLHGDWIRDPATPIVTEFLKRWKSGAVVTLGSGSDLTATTTGKLGIR
EIQWMDVDVGLCDPNLYPNATVPIYTFEEAEIAFGAQVJHPSPMRPARGDIP
VRVKNKSYNKPAGTLITKQREMDKVLTVLSIKSVNMLDIVSRMGJGSGFLAKYFS
IPEDLGIISVDCVATSEVSIIVSLDPSKIMSESLIOELIDHVELEKIAVHLIQQA
IISLIGNRSRLIEKAFQVYRKSGVAVVQMSQASRVNMSLIVHDSBAKQCKIKLH
QAFEDVDVTEVERALHMS"

ORIGIN
Query Match 39.9%; Score 791.4; DB 8; Length 1343;
Best Local Similarity 84.6%; Pred. No. 1.9e-181;
Matches 930; Conservative 0; Mismatches 151; Indels 18; Gaps 3;

QY 629 TTTGAGATGCAATGTCACGAGATTTTCTGCTTATTGGAACAAATTCGTCAG 688
DB 1 TTTGGGAGTGCATGCTACACGAGATTTTCTGCTTATTGGAACAAATTCGTCAG 60
QY 689 GCACGCGAGTATGACGATTTGATTTGCTTATTACACATGATGATTTGTTATCG 748
DB 61 GCACGCGAGTATGACGATTTGATTTGCTTATTACACAGAGATTTGGAACGCG 120
QY 743 GATATCTTGAAGCAACCTATCTCTGCTGTGGAGAGACATTCATGGGACTGATACG 808
DB 121 GACATCTTGAAGCGAGCTTATCTCTGCTGTGGAGAGATTCATGAGACTGATTCGT 180
QY 809 GATCCAGCATACCTGTTGTTACTGGGTTCTTTGGAGAGGCTGGAAATCGTGCTGTA 868
DB 181 GATCCAGCATACCTTATTTGTTACTGGGTTCTTTGGAGAGGCTGGAAATCGTGCTGTA 240
QY 869 ACTACTTTAGCCGAGAGTGTAGTGAATTCATCTCAACACATTTGGTAAAGCTTTGGGA 928
DB 241 ACCACTTTAGCCGAGAGTGTAGTGAATTCATCTCAACACATTTGGTAAAGCTTTGGGG 300
QY 929 CTGAGAGAAATCAGGTATGAGAAAGTGTGATGTTACTTACTTGGATCCAAATATC 988
DB 301 TTGAGAGAAATCAGGTATGAGAAAGTGTGATGTTACTTACTTGGATCCAAATATC 360
QY 989 TACCCATGCAAAAGCTGTTCAATCTTAACATTTGAAGAGCCACGAACTTGCTTAT 1048

DB 361 TACCCGAATGCACACACCTGTGCCATCTTAACTTTGAAGAGCGACGAACTTGCTTAT 420
QY 1049 TTTGGTGTCAAGTTTGTGCATCCACAAATCGATGAGACCTGTAGAGAGTGAATTTCCA 1108
DB 421 TTTGGTGTCCAGGTTTATCATCCACAAATCTATGAGACCTGTAGAGAGTGAATTTCCA 480
QY 1109 GTTAGGTTTGAAGATTCATACACCCCTTAAAGCTCCAGCACCTTATTACAGCAAGA 1168
DB 481 GTCAGGTTTGAAGATTCATACACCCCTTAAAGCTCCAGCACCTTATTACAGCAAGA 540
QY 1169 GACATGATTAAGTCTGTTGTACTAACTGATGTCCTCAAGTCAAAATGTCATATG 1228
DB 541 GAGATGATTA-----GTTGTGTACTAATGCAATGTCCTCAAGTCAAAATGTCATATG 594
QY 1229 TTGACATTTGTAGCACTCGGATGCTTGTGTCAGTATGTTTCTGGCAAGGTTATCAGGT 1288
DB 595 TTGATTAATAGTAAAGCACTCGGATGCTTGTGTCAGTATGTTTCTGGCAAGGTTATCAGGT 654
QY 1289 ATTGCTATATGAAGATCTATGATTAATCTGTGATTTGTGCTACAGTGAAGTTAGT 1348
DB 655 ATAT-----TCGAGATCTTGGGATATCTGTATGATTTGTGCTACTAGTGAAGTTAGT 708
QY 1349 GTTTCGTGTCACTTGATTCATCAAGATCTGAGTGAAGGAACTGATPACAGCGCAAGT 1408
DB 709 ATTTCTGTGTGCTTGTGATTCATCAAGATCTGAGTGAAGGAACTGATPACAGCGCAAGT 762
QY 1409 GAACCTGACATGATGTTGATGAGACCTGAGAAATATACCAATTTGTGCTTACTTACGACAG 1468
DB 763 GAACCTGACATGATGTTGATGAGACCTGAGAAATATACCAATTTGTGCTTACTTACGACAG 822
QY 1469 AGCGGCAATATTTCACTTATGCAAAATGTGAGGCAATGCTCTGATACTAGAAAGACG 1528
DB 823 AGCGGCAATATTTCACTTATGCAAAATGTGAGGCAATGCTCTGATACTAGAAAGACG 882
QY 1529 GCACGTGTGCTGAGGAAAGTGGGTTATGTTGATGATCTTGGCAAGAGCGTCAAAG 1588
DB 883 TTTCAAGTGTGAGGAAAGTGGGTTATGTTGATGATCTTGGCAAGAGCGTCAAAG 942
QY 1589 GTTAACATGTGCTGATGTCATGATGAGCAGTCAAGGCACTGTGAGAGCCCTTCAT 1648
DB 943 GTTAACATGTGCTGATGTCATGATGAGCAGTCAAGGCACTGTGAGAGCCCTTCAT 1002
QY 1649 CAGCGCTTCTTGAAGACGATGCTTATCACAAATGCAAGCGAGAACTTACTGTGGGC 1708
DB 1003 CAGCGCTTCTTGAAGACGATGCTTATCACAAATGCAAGCGAGAACTTACTGTGGGC 1062
QY 1709 TGATCAGGTAGGCTTGGC 1727
DB 1063 TGATCAGGTAGGCTTGGC 1081

RESULT 6
AK121930 1916 bp mRNA linear PLN 29-OCT-2003
LOCUS
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J033106KX2, full
insert sequence.
ACCESSION AK121930
VERSION AK121930.1 GI:37991553
KEYWORDS
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta, Liliopsida; Poales; Poaceae;
Eriarctioideae; Oryzaceae; Oryza.
REFERENCE
1
The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team,
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, Y., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yanae, W., Suzuki, K., Li, C.,
Ohtsuki, K., Shieniki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group, Ohtomo, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE
japonica rice
Science 301 (5631), 376-379 (2003)

22752273
12869764

JOURNAL
REFERENCE
AUTHORS
TITLE

3 (bases 1 to 1916)

Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of
Agricultural Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)

This clone is one of the 32K full-length cDNA clones from japonica
rice.

COMMENT

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Setoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
Ohnada, B., Yanagi, M., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T.,
Yamamoto, M., and Nakahama, Y.

PAIS Genome Sequencing & Analysis Group: Oono, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M.,
Kobayashi, M., Kodama, T., Kuroseki, T., Kusumegi, T., Lu, M.,
Masuda, H., Miura, J., Mizuno, K., Naitaka, R., Nishikawa, J., Oka, M.,
Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsumoda, Y., Ueda, M.,
Xie, Q., Yokomizo, S., Yoshimura, A., Matubara, K., and Murakami, K.

Genome Exploration Research Group in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiwamoto, K.,
Hiraoaka, T., Horii, F., Iida, O., Imamura, K., Imoto, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Nomura, K., Nunnaki, R., Ohta, N.,
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakai, K., Saitoh, N.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M.,

Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaki-Akashira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yaenunishi, A., and Hayashizaki, Y.

FEATURES

source
1. 1916
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultiyar="Japanbare"
/db_xref="taxon:39947"
/clone="J033106K22"

ORIGIN

Query Match 39.7%; Score 787.6; Db 8; Length 1916;
Best Local Similarity 74.0%; Pred. No. 1,7e-180;
Matches 1058; Conservative 0; Mismatches 351; Indels 21; Gaps 4;

274 GCTCAGCGTGTGATGAGATTCCGGGGGCTCTGCTGTGTGCGCCGAGATGGCTCA 333
243 GTTCACCGTCGCCATGAAAGTTCCGGCGGTGTGCTGGCGGTGCGGAGCGATGGGGA 302
334 GGTGGCGGCGCTCATCTCTGACGTTCCCGAGAGCGCCCGTGTCTCTCGCAT 393
303 GGTGGCGATTTCTATCTCAGCTTCCCGAGAGAGATCCCGTGTCTCTCGCAT 362
394 GGGGAAACCAACCAACACCTTCTCTCTGTGAGAGAAAGCGATGAGGTGAGTTAT 453
363 GGGGAAAGCCACCAATTAACCTCTCTCTGCGGAGAGAAAGCGATGAGGTGAGTTAT 422
464 CCAATGTTTCAAAATTCGAAGAGTAATGTCGAAGAAAGCGATGAGGTGAGTTAT 513
423 GAAAGCGGTGAATTTCCAGAGCTCGGAGTTACAGAGCTTCAATTTGACATTTA 482
514 TGAATGGAATTCGAAGATTCGTAATACAGACCTTT--ATGAATGAGCAATAT 570
483 TGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 542
571 GAAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 630
543 TAAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 602
631 TGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 690
603 CGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 662
691 ACGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 750
663 TCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 722
751 TATCTTGAAGCAACCTATCTCTGCTGTGGAAGAGATTCATGAGGAGCTGATACGA 810
723 CATTTTGAAGCAACCTATCTCTGCTGTGGAAGAGATTCATGAGGAGCTGATACGA 782
811 TCCAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 870
783 CCGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 842
871 TACTTGAAGCAACCTATCTCTGCTGTGGAAGAGATTCATGAGGAGCTGATACGA 930
843 CACTTGAAGCAACCTATCTCTGCTGTGGAAGAGATTCATGAGGAGCTGATACGA 902
931 GAGAGAAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 990
903 AAGAGAGATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 962
991 CCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1050
963 TCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1022
1051 TGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1110
1023 TGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1082
1111 TAGGTTAAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1170

Db	1083	TAGAGTTAAGAAATTCATATAACCGTCGAGCACCCTGGTACTGTTATTACTTAACAAGGA	1142
Qy	1171	CATGATTAANGCTCTGTTGTACTACTAGCATAGTCTCAAGTCAATGTCATATGTT	1230
Db	1143	TTTGAGCAGAG-----TATATTAACCGCATTTGATTGAAGTAAACATTAATGCT	1196
Qy	1231	GGACATTGTGAGCACTCGAGTGTGTGTAGATGTTTTCGGCAAGGGTACAGGTAT	1290
Db	1197	GGATATAGGAGACACAGAGATCTGGCCAGATGTTTTCGGCAAGGGTTCCTCAT	1256
Qy	1291	TTGCTATTGTAAGATCTATATATCTGTGATGTTGTCTACCACTGAAGTTAGT	1350
Db	1257	AT-----TTGAAGATTAGATCTCTCTGTTATGTTGGCTACCACTGAAGTCAAT	1310
Qy	1351	TTCTGTGTCACTTGATTCATCAAAAGTCTGAGTGAAGGAAGTATACAGCGCAAGTGA	1410
Db	1311	ATCATTTGACACTAGATTCATCAAAAGTATGTTGTGTGATTTGATTCACGAG-----GA	1364
Qy	1411	ACTTGAACATGTATGTTGAAGAGCTTGAGAAATAGCAATGTTGCTTACTTCAAGCAG	1470
Db	1365	ACTGTATCATGTATGTTGAAGCACTTGAAATTTGCAATTTCTACTGCGACAGCAG	1424
Qy	1471	GGCGATTAATTCACTTATCGGAATGTGAGCAATCGTCTCTGATCTAGAAAAGACGGG	1530
Db	1425	GTCATCATCTCTCATGATGAGGAATGTACAGAGATCGTCATTTGTTGAAAGGCTTT	1484
Qy	1531	ACGTGTCTGAGGAAAAGTGGGGTTAAATGTTAGATGATCTTCGCAAGAGGCTCAAGGT	1590
Db	1485	CAATGTTCTACGGAGAAATGGGGTTAAATGTTAGATGATCTTCACAGGGGATCCAAAGT	1544
Qy	1591	TACATGTGCTGATATGTCATGATAGCGATCAAGGCAAGGCACTCTGAGAGCCCTTCA	1650
Db	1545	GAAACATTTATGTTGTGTCTCCATGATAGCAGGCGAAACAGTATACAGGCCCTTCATTC	1604
Qy	1651	GGGGTCTTTGAAGACGATGTCCTATCACAAGTCGAGCGGAGAACTTAC	1700
Db	1605	AGCATCTTTGAGAACGGTTCTTGTCAAGATGTGAGAGAGGATCTTC	1654
RESULT 7			
AKI02162			
LOCUS			
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:J033086G22, full insert sequence.		
ACCESSION	AKI02162		
VERSION	AKI02162.1 GI:32987371		
KEYWORDS	FLI CDNA; CAP trapper.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.		
REFERENCE			
AUTHORS	1 The Rice Full-length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Onoda, B., Yanagi, M., Sugita, K., Ueda, M., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niihara, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Mura, J., Kusunegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., Aikawa, T., Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oosato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.		
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice		

JOURNAL	Science 301 (5631), 376-379 (2003)
MEDLINE	22752273
PUBMED	12869764
REFERENCE	2 (bases 1 to 1983)
AUTHORS	Adachi, J., Aizawa, K., Aikawa, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashidume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Hotta, I., Iida, J., Iida, J., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kuribara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Mura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niihara, J., Nishii, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Oosato, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.
TITLE	Direct Submission
JOURNAL	Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skk@kenken.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT	This clone is one of the 28k full-length cDNA clones from japonica rice.
URL	http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and Yamamoto, M.	
FAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Mura, J., Mizuno, K., Narikawa, R., Niihara, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Kurakawa, K.	
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Aikawa, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kuribara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Nakamura, M., Nishii, K., Nomura, K., Numasaki, R., Ohno, M., Oosato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A., and Hayashizaki, Y.	
Location/Qualifiers	
1..1983	
/organism="Oryza sativa (japonica cultivar-group)"	
/mol_type="mRNA"	
/cultiivar="Nipbare"	
/db_xref="taxon:39947"	
/clone="J033086G22"	
ORIGIN	
Query Match	39.2%; Score 777.6; DB 8; Length 1983;
Best Local Similarity	73.9%; Pred. No. 4.6e-178;
Matches 1048; Conservative	0; Mismatches 350; Indels 21; Gaps 4;
Qy	273 AGCTCAGCGTGTGATGAAGTTGGGGGGGCTCTGCTGCTGCTGCGCCGAGAGATGAGCTG 332

D	320	AGCTACCGTGGTGTGTAAGTTCCGGCGGGTGTCAGTGCGCTCGCGCGAAGCGAG	379
Q	333	AGGTGGCGGGCTCATCTGACGTTCCCCGAGAGCGCCCCCGTCGTCCTCTCGCCA	392
D	380	AGGTGGCCCATCTCATCTCGACGTTCCCGGAGGCGCGGTGATGTGATCTCTCGCCA	439
Q	393	TGGGGAAAACCAACAACACTTCTCCTGCTGGAGAGAGCGAGTGTGAGATTA	452
D	440	TGGGAGAGACCAACAACAGCTTCGTGATGCTGGGAGAAAGCGGTGTGGTGCA	499
Q	453	TCCATTTTCTGAATCGAAGTGGAAATATGTCTAAAGCTTACATACAAAGCGTGG	512
D	500	CAAAATGTGTGGAGCTTATAGAGCTCACCTTTATTAAGAAATTCATTTGGACCAATG	559
Q	513	ATGAATCTGAGCTTCCAAAGNATCTGTATAACAGCTTT--ATGAATCGAGCAACTAT	569
D	560	ATCAACTTGATTTGATAGTCAATTTGTTCCGGCTTGTGAGATGAGCTTGACCACTTC	619
Q	570	TGAAGATATCGCTATGATGAAAGAGCTGACGCTTAGACAGTACTCTTTGTTTAT	629
D	620	TTAAGGAATTTCTATGATGAGAGGCTTACTTTAGACAAAGACTATCTTGTTCCT	679
Q	630	TTGGAATATGACATGTCCACAGGATTTTTCGCTTATTTGAACAAAATTCGTCAAG	689
D	680	TGGTAAATGACATGTCCAAAGAAATTTGGCGCACTTTGAATAACTTGGGTGAAG	739
Q	690	CACGCGATATGACGATTTGTATATGGTTTCATTAACATGATGAATTTGGTATCGG	749
D	740	CACGTAGATATGACATTTGAAATAGGCTTTATMACACTGATGATTCACAACAGCTG	799
Q	750	ATATCTTAAGAACCACTATCCCTGCTGTGGAGAACTTCATGGGGACTGGATACAG	809
D	800	ATATCTGGAACCACTTATCTCGCATTTGAAAGAGCTTACATATGGGAATGGGTACTG	859
Q	810	ATTCAGCATACCTGTTGTACTTGGGTTCTTGGGAAAGGCTGGAATCTGGTCTGTAA	869
D	860	GTCTCTCTATTCOAATAGCACTGGTTTTCGCGAAAGGAGTGAATCTGGTGCATCA	919
Q	870	CTACTTTAAGCCGAGGTGGTAGTGACTGACCTGTACAAACATTTGGTAAAGCCTTGGAC	929
D	920	CCACTTTAAGCAGGGGTGGTAGTGACCTGACAGCTATACATTTGGCAAAAGCTTGGAT	979
Q	930	TGAGAAATTCAGGATGAGAAAGATGTTGATGGTGTACTTCTGTGATCCAAATATC	989
D	980	TAAAGAAATTCAGGTTTGGAAAGATGTTGATGGTGTACTTCTGTGATCCAAATATC	1039
Q	990	ACCCACATCCAAAGACTGTTCCACTTACATTTGAAGAGGCGACAGAACTTGCTTAT	1045
D	1040	ATCCAAATCGAAACCTGTGCCATCTTGACATTTGATGAGGCGAGCACTTGCTTAT	1099
Q	1050	TTGGTCTCAGGTTTGGATTCACAAATGATGAACCTGCTAAGAAAGGTGATATCCAG	1105
D	1100	TTGGTCCAGAGTTTGGATTCACCAATGATGACACAGCTAGAGAAAGGTGATATCCG	1155
Q	1110	TTAGGGTTAAGATTCATACAAACCTTAAGCTCCAGGCACTTATTAACAGCAAAAG	1165
D	1160	TCAGGGTTAAACCTTATTAACCGCGAGCCCTGTGACATCATTAACAAAGTAAAG	1215
Q	1170	ACATGGAATAAAGTCTGTGTGTACTAACTACAGATAGTGTCAAGTGAATATGCAATAT	1225
D	1220	ATATGAGTAA-----GACTGTTTGGACTAGTATTTGACTTAATATCAATATTAACATG	1275
Q	1230	TGGACATTTGTAGCACTGGATGCTTGTGTGAGATGGTTTTCGGCAAGGATATCAGTAA	1285
D	1274	TTGATTTAGTAGACACACGAGTGTGGACAATATGGTTTTCAGTAAAGTCT-----	1325
Q	1290	TTTGTGATATTTGAAGTCTATGATATCTGTGAGATGGTGTGCTACAGGAAAGTATAGT	1345
D	1328	TTTCAATTTTGAAGTTTGGGCACTCTGTGATTTGTAGCTTACAGTGAAGTACGCA	1385
Q	1350	TTTCTGTGTCACTTGAATCCAAAGATCTGAGTAGGGAATGATACAGCAGCAAGTG	1405

[illegible]

QY 284 GTGATGAAAGTTCGGGGGGTCTCGGTGTGTGGCCGCGAGATGGCTGAGTGGCCGGC 343
 Db 259 GTGATGAAAGTTCGGGGGGTCTCGGTGTGTGGCCGCGAGATGGCTGAGTGGCCGGC 318
 QY 344 CTGATCTGACGTTCCCGAGAGAGCCCGCGTGTGTGTCTCTCGCATGAGGAGAAAC 403
 Db 319 CTGATGATGAGTTTCCCGAGAGAGCCCGATGTGTGTCTCTCGCATGAGGAGAAAC 378
 QY 404 ACCAACAACCTTCTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 463
 Db 379 ACAAACAAGCTTTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 438
 QY 464 GAATCGAAGAGTGAATATGTCACAAAGCTTATCATCAAGCGGTGGATGAATTTGA 523
 Db 439 AGATGAGAGAGTGTGTGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 498
 QY 524 CTTCAGAGATCTGTATACAA---GCCCTTATGAGCTGAGAGAGAGAGAGAGAGAG 580
 Db 499 GTGAGAGAGATCTGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 558
 QY 581 GCTATGATGAAGAGCTGACGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 640
 Db 559 GCTATGATGAAG 618
 QY 641 ATGTCCACAGAGATTTTCTGTGTATTTGAACAAATTCGTGTCAAGAGAGAGAGAG 700
 Db 619 ATGTCCACAGAGATTTTCTGTGTATTTGAACAAATTCGTGTCAAGAGAGAGAGAG 678
 QY 701 GAGCATTTGATTTGTTTCATTAACATGATGATTTGATTTGATTTGATTTGATTTG 760
 Db 679 GATGCAATTTGATTTGTTTCATTAACATGATGATTTGATTTGATTTGATTTGATTT 738
 QY 761 GCAACCTATCTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 820
 Db 739 GCAACTATCTGAGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 798
 QY 821 CCGTGTGATGAGGGTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 880
 Db 799 GCAATGTTGATGAGGGTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 858
 QY 881 CGAGTGTGATGAGCTGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 940
 Db 859 AGAGGGGAGATGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 918
 QY 941 CAGGTATGAGAGAGATTTGATGATGATTTGATGATTTGATGATTTGATGATTTGAT 1000
 Db 919 CAGGTATGAGAGAGATTTGATGATGATTTGATGATTTGATGATTTGATGATTTGAT 978
 QY 1001 AAGACTGTTCCATTAACATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1060
 Db 979 GAACCTGTTCCATTAACATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1038
 QY 1061 GTTTTGATTCACAAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1120
 Db 1039 GTTCTACATCCAAAGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1098
 QY 1121 AATTCAATCAACCTTAAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1180
 Db 1099 AATTCTTACAACTTAAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1157
 QY 1181 GGTGTGTTGATTAACATGATGATGATGATGATGATGATGATGATGATGATGATG 1240
 Db 1158 -----GGAGTATTAACAGCATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1212
 QY 1241 AGCACTCGAGTCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1300
 Db 1213 AGCACTCGAGTCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1266
 QY 1301 GAAGATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1360
 Db 1267 GAAGATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1326

QY 1361 CTTGATCCATCAAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1420
 Db 1327 CTGATCCATCAAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1386
 QY 1421 GTATGAGAGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1480
 Db 1387 GTTGTGAAG 1446
 QY 1481 TCACCTATCGAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1540
 Db 1447 TCTCTATTTGAAG 1506
 QY 1541 AGGAG 1600
 Db 1507 CGAAG 1566
 QY 1601 CTGATGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1660
 Db 1567 TTGTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1626
 QY 1661 GAAGAGAGATGTC 1672
 Db 1627 GAGAGTGAAGCTC 1638

RESULT 9
 AX507049
 LOCUS AX507049 1710 bp DNA linear PAT 27-SEP-2002
 DEFINITION Sequence 1744 from Patent WO0216655.
 ACCESSION AX507049
 VERSION AX507049.1 GI:23388286
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
 AUTHORS Harper, J.F., Krieps, J., Wang, X. and Zhu, T.
 TITLE Stress-regulated genes of plants, transgenic plants containing
 same, and methods of use
 JOURNAL Patent: WO 0216655-A 1744 28-FEB-2002;
 The Scripps Research Institute (US) ; Syngenta Participations AG
 (CH)

FEATURES
 source Location/Qualifiers
 1..1710
 /organism="Arabidopsis thaliana"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3702"

ORIGIN
 Query Match 35.6%; Score 707.6; DB 6; Length 1710;
 Best Local Similarity 70.5%; Pred. No. 4.9e-161;
 Matches 1008; Conservative 0; Mismatches 401; Indels 21; Gaps 4;

QY 254 ATGAGAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAGAGAGAT 313
 Db 235 ACGAGAGATGATGAGAGAGATGATGAGAGATGATGAGAGATGATGAGAGATGATGAG 294
 QY 314 TGGCGCGAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAGAGAG 373
 Db 295 TCAAGTGAAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAGAGAG 354
 QY 374 GTGCTGTCTCTGTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 433
 Db 355 GTCATGTTCTCTGTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 414
 QY 434 GCAGTGAAGTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 493
 Db 415 GCGGTGATGTTGAGTGTCTTAAGATGATGATGATGATGATGATGATGATGATGATGAT 474
 QY 494 CTACATATCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 550

Db 475 TTGCATATCAGAGCGGTGAAAAGAGCTCAACATTGATCCTCTGTATTTTGGACCTATTTG 534
 Qy 551 TATGAACTGAGAGCAACTATTTGAAGGTATCGCTATATGAAGAAGCTGACGCTTAGAGCC 610
 Db 535 GAGGAACTGAGAGCAACTCTTGAAGGCAATGCCATGATGAAGAAATTTGACACTTGCAGCC 594
 Qy 611 AGTGAATACCTGTTGTTCTATTGGAAGATGATGCCACAGATTTTTTCTGCTTATTTG 670
 Db 595 AGAGATTAATCTAGTCTCTTTTGGAGAGTGTTCCTCAAGAGATTTTTGCTGCTTATCTT 654
 Qy 671 AAAAAAATTCGTGCTCAAGAGCAAGAGATGAGCACTTTGATATGTTGTTTCAATACACT 730
 Db 655 AATACAAATCGGTTCAGAGCAAGCAATATATGATCAATTTGAATGGTTTCAATACAG 714
 Qy 731 GATGAATTTGGTAAATGCGGATATCTTGAAGCAACCTATCTCTGTTGGAGAGACTT 790
 Db 715 GATGATTTCACAAAATGGGGATATCTGGAGGCACTTATCCAGCTGTGGCAAGATTTA 774
 Qy 791 CATGGGAGCTGGATACAGAGATCCAGCATCTGTTGTTCTGGGTTCTTGGAGAGGGC 850
 Db 775 TATGATGTTGGATGCAATGATCTGCTGTTCTTATTTGAAGGTTTCTTGGAGAGGGT 834
 Qy 851 TGAAGATCTGCTGCTGATCTTATAGCCGAGGTGTAGTGACTTGAATGCTTACACCC 910
 Db 835 TGAAGAACTGGTGGCGTTACTACTTGGGTAGGGGTGGAGTGAATTTGACGCAACACA 894
 Qy 911 ATTGTTAAGCTTGGAGCTGAGAGAAATCAGGTATGAGAAAGTGTGATGGTGTACTT 970
 Db 895 ATTGTTAAGCGTTGGGTTTGAAGAGATTTAGGTGTGAGAAAGTGTGATGGTGTCTTA 954
 Qy 971 ACTTGTGATCCAAATATCTTACCAATGCAAGAGCTGTCCATTAATTAATTTGAAGAG 1030
 Db 955 ACATGTGACCCCTACTATATATAGAGGCTACACAGATACCAATCTGACATTTGACGA 1014
 Qy 1031 GCCACAGACTTGTCTTATTTTGGTCTCAGTTTGTGATCCAAATGATGAGAACTGCT 1090
 Db 1015 GCAGCGGAGCTAGCTTATTTTGGTGCACAGGTCTTGCACCCAGCTGAAGAGACAGCA 1074
 Qy 1091 AGAGAGGTGATATTCAGTTAGGGTTAAGATTCATACCAACCTTAAGCTCCAGGACCC 1150
 Db 1075 AGAGAGGTGATATTCCTGTTAGGGTTAAATTTCTATACCTTAAGCTCTCTGAGACT 1134
 Qy 1151 CTATATACCAAGCAAGAGCATGATTAAGGTCTGTTGATCTATACATAGATGCTC 1210
 Db 1135 ATCATCACTTAAACCAAGAGCATGATCAACAGAG-----TATCTTAAGACACTTGTCTG 1188
 Qy 1211 AAGTCAATGTCACTATGTTGAGATTTGAGCACTCGAGTGTGCTGATGATGTTT 1270
 Db 1189 AAGCTATGTGACCACTGCTGATATGAGCAAGCAGCGAATGCTTGGTCAAGTTGGCTTT 1248
 Qy 1271 CTGCGAAGGTATCAGTATTTGCTATATGAGATCTTATGATATCTGATGTTGTT 1330
 Db 1249 CTGCGAAGGTATTTCTCATAT-----TTGAGGAGCTTGGCAATTCGATGATGTTGTT 1302
 Qy 1331 GCTACCAAGTAAAGTATGTTCTGCTGCACTTGTATCATCAAGAGCTGAGATAGAGAA 1390
 Db 1303 GCCACTATGTAAGTCAATATCTCTGACACTGATCTCTCAAACTTTGAGACAGAGAA 1362
 Qy 1391 CTGATACAGCAGGCAAGTGAATTGACCATGTAGTTGAAGAGCTTGAAGAAATGACAT 1450
 Db 1363 CTGATTTCAACAG-----GAGCTGATCATGTATGTAAGAACTGAGAAATTTGACAGTT 1416
 Qy 1451 GTTGCTTACTTTCAGCAGAGGCGCATATTTCACTTAATGGAAATGTGAGCAATGCTCT 1510
 Db 1417 GTGATCTTTTAAAGAGAGCAATCATCTCTTAATTTGGAAATGTTCAACATTTCTCTC 1476
 Qy 1511 CTGATACAGAAAGACGAGAGCTGTGCTGAGAGAAAGTGGGTTAAATGTTCAAGATATC 1570
 Db 1477 CTGATTTAGAGAGGCGCTTCAATGTTCTTATACCAAGGTGTCATATGTCAGATGATA 1536
 Qy 1571 TCGCAAGAGAGCTCAAAAGTTAATGATGCTGATATCCATGATAGCGATGCAAGAGCA 1630
 Db 1537 TCACAAGAGATCCAAAGTTAAACATTTCTTTATAGTAAACGAAGCTGAAGCTGAAGGT 1596

Qy 1631 CTCGTAGAGACCCCTTATCATAGGCGTTCTTTGAAGAGCATGCTCTATACCA 1680
 Db 1597 TGTGTTAGGCTCTTTCACAAATCTTCTTGAGAGCGGTGACCTCTCAGA 1646
 RESULT 10
 BT000493
 LOCUS
 DEFINITION
 Arabidopsis thaliana aspartate kinase (At5g13280/T31B5_100) mRNA, complete cds.
 ACCESSION
 BT000493
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1710)
 REFERENCE
 AUTHORS
 Cheuk, R., Chen, H., Kim, C.J., Shim, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y., Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Tortum, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinzaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
 Arabidopsis ORF clones
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 Cheuk, R., Chen, H., Kim, C.J., Shim, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y., Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Tortum, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinzaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
 Direct Submission
 Submitted (25-SEP-2002) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 COMMENT
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RFL CDNA (RFL CDNA: 'RIKEN Arabidopsis Full-length CDNA') : Seki, M., Narusaka, M., Ishida, J., Hayashizaki, Y., and Shinzaki, K.
 The Salk, Stanford, PEGC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RFL CDNA: Cheuk, R., Chen, H., Kim, C.J., Shim, P., Baw, P., Baw, P., Chan, M.M., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Oodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Tortum, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Davis, R.W., Theologis, A., and Ecker, J.R.
 Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinzaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.
 FEATURES
 source
 1..1710
 location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="U14700"
 /ecotype="Columbia"
 /note="This clone is in PUNI 51"
 1..1710
 /note="aspartate kinase"

CDS

```

/codon_start=1
/product="At5g13280/131B5_100"
/protein_id="AA18062.1"
/db_xref="GI:23308185"
/translation="MAATRVCCSHNAFRLPLTRHNSPTLPISLRVDPPTLKL
SLPIDGSSIRKVSQSRNIVRAVLEKTEKTEATTEDEKGTICVMFGSSVSAER
MEVADILITPEPESPVAVLSMGKTTNLLAGEKAVSCVNSAIEELSLIKELH
IRVKEINIDPSVILITLLELBOLKIGLAMKELTLETRDVLSPFESLSTIFAAYL
NRIKVRQXOYPAFEIGPTTDDPTNGDILKATVPYAKRPYDMDHPDAPITVTEGL
KMKRTGAVTTLRGSSDITATTIGKALGKELQVAKYVDGLTLDPTTYKRAIPVPL
TFDEAEILAVFGAQLVHPQSRPABEGLPVKNSINPAPGTTIKTIDMKSLIT
SIVLRNVTMLDIASTRMLQGVGLAKVFSIFEELGISVVAATSEVISLTLDPSTL
WSRELIOQLDHVLEBKIAVNLKRAIISLIGNVCHSLLEBAFVLYTKGVN
VOMISOGASKVNIISPIVNEAEKGCVALHKSFPESGLSELILQPLNGNSPRTLO
VEN"

```

ORIGIN

```

Query Match      35.6%; Score 707.6; DB 8; Length 1710;
Best Local Similarity 70.5%; Pred. No. 4,9e-161;
Matches 1008; Conservative 0; Mismatches 401; Indels 21; Gaps 4;

QY 254 ATGAGAGGATGAGGGGATCAGCTCAGCCGTGATGAAGTCCGGGGGCTCTCGGTGCG 313
DB 235 ACGGAGGTGATGAGAGGATATCAGTGCCTGATGAGAGTTGGTGAATCTCGGTGCG 294
QY 314 TCGCCGCGAGAGTGGCTGAGGTGCGCGCCCTCATCTGACGTTCCCGAGAGCGCC 373
DB 295 TCAGCTGAGAGATGAGAGAGTGTCTGATTTGATTTTCCGGAAGAAAGTCCC 354
QY 374 GTCGTCCTCTCTCGCCATGGGGAAAACCCCAACACCTTCTCTGCTGAGAGAG 433
DB 355 GTCATGTTCTCTCTGCTAAGGGAAAACCAACCAATCTCTGCTGCGGAGAGAG 414
QY 434 GCAGTAGGCTGAGATTATCCATGTTCTGMAATCGAAGATGGAATATGTCMAAGC 493
DB 415 GCGGTTAGTTGGTGGTTTCTAATGATCTGAGATTGAGAGATTGACATTAAGAA 474
QY 494 CTACATATCAAGCGGTGATGAATCTTGACT--TCCAGNATCTGMAATCAAGCTT 550
DB 475 TTGCATATGAGACGGTGAAGAGCTCAACATGATCCCTGTTATTTGACCTATT 534
QY 551 TATGAAGTGAAGCAATATTGAAAGTATCGCTATGATGAAGAAAGCTGACCCTAG 610
DB 535 GAGAGATGAGACCACTCTCTGAAGAGCATGACCATGAGAGAAATGACACTCGAA 594
QY 611 AGTACTACCTGTTTCAATTTGAGAGATGATGACACAGAGATTTTCTGCTTATT 670
DB 595 AGGATTTACTAGTCTCTTTTGAAGAGTGTGTTGTCACAGAGATTTTGTGCTTAT 654
QY 671 AACAAATTCGTGTCAAGGACGAGATGACGATTTGATGATGATGTTTATTAACA 730
DB 655 AATACATCGGTGTCAAGGACGCAATATGATGATTTGAATTTGTTTATTAACA 714
QY 731 GATGAATTTGATGATGATGATTTGAGAGCAACTTATCTGCTGTTGCAAGAGCT 790
DB 715 GATGATTTCACAAAATGGGATATCTCGAGGCAACTTATCCAGCTGTTGCCAAGA 774
QY 791 CATTGGGACTGATTCAGATTCAGAGNACCTGTTGTTCTGCGGTTCTTGGGAAG 850
DB 775 TATGATGATTTGATGATGATCTCTGCTGTTCTTATTTGAACAGGTTCTTGG 834
QY 851 TGGAAATCTGTGCTGTAACTATTAGCCGAGGTGATGATGATGATGATGATGAT 910
DB 835 TGGAAATCTGTGCTGTAACTATTAGCCGAGGTGATGATGATGATGATGATGAT 894
QY 911 ATTGTAAGCTTGGGACTGAGAGAAATTCAGATGATGATGATGATGATGATGAT 970
DB 895 ATTGTAAGCTTGGGACTGAGAGAAATTCAGATGATGATGATGATGATGATGAT 954
QY 971 ACTTGTATCAAAATATCTTACCAACATGCAAGAGCTGTTCCATCTTAACATTTGA 1030
DB 955 AACTGTGACCCCTATATATTAAGGAGCTACACAGATGACATCTGACATTTGACAA 1014

```

```

QY 1031 GCCACAGAACTGTGATATTTTGTGTGTGCTCAGGTTTGCATCCACATGATGACCTGCT 1090
DB 1015 GCAGCCGAGCTAGCTTATTTTGTGTGTGCTCAGGTTTGCATCCACATGATGACCTGCT 1074
QY 1091 AGAGAGGTGATATTTCCAGTTAGGTTTGAAGATTCATACCACTTAAAGCTCCAGGAC 1150
DB 1075 AGAGAGGTGATATTTCCAGTTAGGTTTGAAGATTCATACCACTTAAAGCTCCAGGAC 1134
QY 1151 CTATTAACAACAAGAGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1210
DB 1135 ATCATCACTTAAACAAGAGACATGATGATGATGATGATGATGATGATGATGATGAT 1188
QY 1211 AAGTCMAATGTCATATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1270
DB 1189 AAGTCMAATGTCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1248
QY 1271 CTGCGAAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1330
DB 1249 CTGCGAAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1302
QY 1331 GCTACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1390
DB 1303 GCTACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1362
QY 1391 CTGATACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1450
DB 1363 CTGATACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1416
QY 1451 GTTGTCTTACTTCAAGACAGAGGCGATATTTCACTTATCCGAAATGAGACATGCT 1510
DB 1417 GTTGTCTTACTTCAAGACAGAGGCGATATTTCACTTATCCGAAATGAGACATGCT 1476
QY 1511 CTGATACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1570
DB 1477 CTGATACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1536
QY 1571 TCGCAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1630
DB 1537 TCGCAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1596
QY 1631 CTGATGAGAGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
DB 1597 TGTGTTCAAGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1646

```

```

RESULT 11
LOCUS AY057674 1971 bp mRNA linear PLN 22-OCT-2001
DEFINITION Arabidopsis thaliana AT5g13280/131B5_100 mRNA, complete cds.
ACCESSION AY057674
VERSION AY057674.1 GI:16323140
KEYWORDS FLI CDNA
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustoside II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1971)
REFERENCE 1
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bower,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,J., Kamuya,A., Karlín-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinzaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
TITLE Arabidopsis CDNA clones
JOURNAL unpublished
REFERENCE 2 (bases 1 to 1971)
AUTHORS Bower,L., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bower,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,J., Kamuya,A., Karlín-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,

```

TITLE
JOURNAL

Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinzaki, K., Davis, R.W., Theologis, A. and Becker, J.R.
Direct Submission
Submitted (24-SEP-2001) Salk Institute Genomic Analysis Laboratory (Signal), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFI cDNAs (RAFI cDNA: "RIKEN Arabidopsis Full-Length cDNA"): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamuya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinzaki, K.

COMMENT

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFI cDNAs: Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Shim, P., Ban, J., Bowser, L., Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Becker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinzaki, K. (RIKEN GSC) and Becker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source

Location/Qualifiers
1..1971
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="RAF09-49-022(R14700)"
/ecotype="Columbia"
1..78
79..1788
/note="aspartate kinase"
/codon_start=1
/product="AT5G13280/T3B5_100"
/protein_id="At1G3305.1"
/db_xref="gi:16323141"
/translation="MAATRVRCCHSNAAFTRLPLTRHNSPTLISINRVDFTLKKL
SLPIDGSSIRKVSGRNIVRAVLEKTEATTEVDEKQITCWMFGSSVASER
MEVADLLITPEESPVILSAMKGTINLLAGEKAVSCVNSASIEELSIKELH
ITVELINIDPEVILVLELEBOLIKGAMKELTLPRLVDFGCESTR.FAAVL
NTIGKAROYDAFEIGFTTDDPTNGDILEATYPAVKRLDDMDHPAVPIVDFLG
KGMKGAATTTGGGSDLTATTTGALGKELIOWKHVDVLTCDPPIVREAPVILV
TTDEAELATYGAOYLHPSNRPAREGSIPIRYKNSINPRAPGIIIRKIDMTKSLIT
STLVKNAVTLMDIATRWLGGVGLAKVESLFEELGVSVDVATSEVSLITDPSKL
WSRELIOQLHVALELEKIAVNLKRAIISLIGNVGRSLLEAFVLTGKN
VOMISOGSKVNIPIVNEABEGCVQALHRSFESDLELLIQPLNGSEVRLQ
VEN"
1789..1971

3'UTR
ORIGIN

Query Match 35.6%; Score 707.6; DB 8; Length 1971;
Best Local Similarity 70.5%; Pred. No. 5e-161;
Matches 1008; Conservative 0; Mismatches 401; Indels 21; Gaps 4;

QY 254 ATGGAGGAGTGGGGATCAGCTCAGCGTGGATGAAGTTCGGGGGGTCTCGGTG 313
DB 313 ACGGAGTGAATGAAGGATATCAGTGGTGAAGTTCGGTTCGTTGGTGGCG 372
QY 314 TGGCCGCGAGATGAGTGGCGCGGCTCATCTGACGTTCCCGAGAGCGCC 373
DB 373 TAGCTGAGAGATGAAGAGAGTGTGATTTGATTTTCCGGAAGAAAGTCCC 432
QY 374 GTGGTGGTCTCTGCTGCTGAGGAAACACCAACAACCTTCTGCTGAGAGAG 433
DB 433 GTGATGTTCTCTCTGCTATGGGAAACCTACCAACAATCTTCTGCTGCGGAGAG 492
QY 434 GCAGTAGGCTGGAGATTAATCAATGTTCTGTAATGAAGAGTGAATATGCTCA 493

DB 493 GCGGTATGTTGTGTTGTTCTTAATGATCGATGAGATGAGAGTTGACATTAAAGAA 552
QY 494 CTACATATCAAGACGGTGAATGAATGAGCT---TCCAAGNATCTGTATACAGCCTT 550
DB 553 TTGATATCAAGACGGTGAAGAGCTCAACATGATCCCTCGTTATTTGACCAATTTG 612
QY 551 TATGAATGAGCAACTTGAAGAGTATCGTATATGAAGAGCTGACCGCTTGAGCC 610
DB 613 GAGAACCTGAGCAACTCCGAAAGGATGCGCATATGAAGAAATGACATTGCAAC 672
QY 611 AGTACATCACTGTTTCAATTTGGAGATGATGCAATGCAACGAGATTTTCTGCTTATTTG 670
DB 673 AGAGTTTCTAGTCTCTTTTGGAGAGTGTTCCTACAGAGATTTTTCCTGCTTATCTT 732
QY 671 AACAAATTCGTGTCAAGGACGCGAGATGACGATTTGATTTGTTTCAATACACT 730
DB 733 AATCAATCGGTGTCAAGGACGCGCAATATGATGATTTGAAATGTTTCAATACAG 792
QY 731 GATGAATTTGTTATGCGGATCTTGAAGCAACCATGCTGCTGCGAGAGACTT 790
DB 793 GATGATTTACAAATGGGAGATTCCTGAGAGCACTTATCCAGCTTTGCGAAGATTA 852
QY 791 CATGGGACTGATACAGATCCAGCGATACCTGTTGTTACTGGGTTCTTGGGAAGGCG 850
DB 853 TATGATGATGATGATGATGATTCCTGTTCTATTGTAACAGGTTTCTTGGGAAGGCT 912
QY 851 TGAATATTTGTTGTTAACTTCTTGAAGGAGTGTGATGATGATGATGATGATGAT 910
DB 913 TGAATAATGTTGTTGTTAACTTCTTGAAGGAGTGTGATGATGATGATGATGATGAT 972
QY 911 ATTGGTAAAGCTTGGGACTGAGAGAAATTCAGGTATGAGAAATGTTGATGATGAT 970
DB 973 ATTGGTAAAGCTTGGGACTTGAAGAGATTCAGGTATGAGAAATGTTGATGATGAT 1032
QY 971 ACTGTGATCCAAATATCTACCAATGACATGACAAAGCTTTCATCTTAATTTGAAG 1030
DB 1033 AATGTGATCCCTACTATATATTAAGCGGCTACACCGATACCAATGATGATGATGAT 1092
QY 1031 GCCACAGACTTGTGTTATTTTGGTGTGAGTTTGGATGATGATGATGATGATGATGAT 1090
DB 1093 GCAAGCAGACTGATTTATTTTGGTGTGAGTTTGGATGATGATGATGATGATGATGAT 1152
QY 1091 AGAGAGGTGATTTCAATTTAGGTTAAGATTTCAATCAACCTTAAAGTCCAGGAC 1150
DB 1153 AGAGAGGTGATTTCTCTGTTAGGTTAAGATTTCAATCAACCTTAAAGTCCAGGAC 1212
QY 1151 CTATTACCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1210
DB 1213 ATCATCTACTAAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1266
QY 1211 AAGTCAATGTTCAATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAG 1270
DB 1267 AAAGCTATGTTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAG 1326
QY 1271 CTGCAAGGATGATGATTTGCTATATTTGAGATTTGAGATTTGAGATTTGAGATTTGAG 1330
DB 1327 CTGCAAGGATTTGCTATTTT---TTGAGAGCTTGGCATTTCCGATGATGATTT 1380
QY 1331 GCTACAGTGAAGTATGTTTCTGTTCTGATCTGATGATGATGATGATGATGATGATGAT 1390
DB 1381 GCTACAGTGAAGTATGTTTCTGTTCTGATCTGATGATGATGATGATGATGATGATGAT 1440
QY 1391 CTGATACAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1450
DB 1441 CTGATTAACAG-----GAGCTGATCAATGATGATGATGATGATGATGATGATGAT 1494
QY 1451 GTTGTCTACTTACAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1510
DB 1495 GTGATCTTTAAAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1554
QY 1511 CTGATCTTAAG 1570

Db 1555 CTGATTTTAGAGAGGGCGTTTCATGTCCTTTATACCAAGGTGCAATGTCAGATGATA 1614

Qy 1571 TCGCAAGGAGCGTCAAGGTATACATGTCGATAGTCCATGATAGCGTCAAGGCA 1630

Db 1615 TCACAAGGAGCATCCAGGTAAACATTTCTTTATAGTAAAGCAAGCTGAAGGT 1674

Qy 1631 CTCGTAGAGCCCTTCATCAGCGCTTTCTTTGAAGCATGTCCTATCA 1680

Db 1675 TGTGTCAGGCTCTTCACAATCTCTTCGAGAGCGGTGACCTCTCAGA 1724

RESULT 12

ATAKLYSI 1980 mRNA linear PLN 02-JUL-1997

LOCUS A.thaliana mRNA for aspartate kinase.

DEFINITION X98873

ACCESSION X98873.1 GI:2243115

VERSION ak-lys1 gene; aspartate kinase.

KEYWORDS Arabidopsis thaliana (thale cress)

SOURCE Arabidopsis thaliana

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1

AUTHORS Frankard, V., Vaucherin, M. and Jacobs, M.

TITLE Molecular characterization of an Arabidopsis thaliana cDNA coding for a nonfunctional aspartate kinase

JOURNAL Plant Mol. Biol. 34 (2), 233-242 (1997)

MEDLINE 97351557

PUBMED 9207839

REFERENCE 2

AUTHORS Frankard, V.M.S.

TITLE Direct Submission

JOURNAL Submitted (02-JUL-1997) V.M.S. Frankard, Institute for Molecular Biology, Lab of Plant Genetics, Free University of Brussels.

REMARK Peptidestrat 65, 1640 Sint Genesius Rode, BELGIUM

REMARK revised by [3]

REMARK 3 (bases 1 to 1980)

REMARK Frankard, V.M.S.

REFERENCE Direct Submission

JOURNAL Submitted (02-JUL-1997) V.M.S. Frankard, Institute for Molecular Biology, Lab of Plant Genetics, Free University of Brussels.

REMARK Peptidestrat 65, 1640 Sint Genesius Rode, BELGIUM

COMMENT Related sequence: X99107.

FEATURES

source

1..1980

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/note="ecotype Columbia"

1..1980

/gene="ak-lys1"

79..1788

/gene="ak-lys1"

/EC_number="2.7.2.4"

/codon_start=1

/product="aspartate kinase"

/protein_id="CAA67376.1"

/db_xref="GI:4376158"

/db_xref="GOA:O23152"

/db_xref="SPRPMEM:O23152"

/translation="MAATRVRCCHSNAAPFRLPRLRNSPTLPSLNRYDPFLKLTSLPIGSSIRKVSIGSGSRNIVRAVLEKKEATVEDEKGTVMKFGSSVASAKRKEVADLITPEESPVIYLSAAGKTNLLAGEKAVSCVSAASEIBELSLIKEDLIRTVKELNDPSVILTYLEBELQILKGLAMMKELTLRFDIVSFGSCLSRIFPAVLNTIGVARYDAFEIGFITTDPTNGDILAEATPAVARLDVDMHDPAYIVTGFGLKGMKTGAVTTLGRGSDLTATIGKALKEIQWKVDVGLTCDPTIKYKATAPVLTTPDEAELAYFGAOVYLHPOSNRPARGEIPYRVKSGSYNPKAGTIIKTRMTKSLITSTIKRNTMTLIDIASTRMTCGOYGLAKYPSIPEEGISVDVATSEUSISLTDPSLTWSRELQDELDRVBELEKIAVYVNLKGRALISLIGNVQHSLLLEAPHLITKGINVOMISQASAKNVSIFVNEABEGCVQALHRSFFESGDLSELLIQPRGNSPVRLIQVBD"

79..336

mat_peptide

337..1785

/gene="ak-lys1"

/product="aspartate kinase"

/EC_number="2.7.2.4"

339..360

/gene="ak-lys1"

/note="conserved aspartate kinase amino acid box, KFGG"

1970

polyA_site

ORIGIN

Query Match 35.6%; Score 706; DB 8; Length 1980;

Best Local Similarity 70.4%; Fred. No. 1.2e-160; Indels 21; Gaps 4;

Matches 1007; Conservative 0; Mismatches 402;

Qy 254 ATGAGAGATTGGGGATCAGCTCAGCGTGGTGAAGTTCGGGGGTCCTCGTGTG 313

Db 313 ACGGAGGTAAATGAAAGGATCACGTCGTGATGAAGTTGGTGTGATCTTCGGTGGC 372

Qy 314 TCGCCCGAGAGATGCTGAGGTGGCCGCTCATCTGACGTTCCCGAGAGCGCCC 373

Db 373 TCAAGTGAAGATGAAGAAAGTGTGATTTGACTTTCCGAAAGAAAGTCCC 432

Qy 374 GTGATGTTCTCTGCGCATGGGAAAAACCAACAACCTTCTCTTGTGAGAGAAG 433

Db 433 GTCAATTTCTCTCTGCTATGGGAAAACTCAACAAATCTTGTGCGGAGAGAAG 492

Qy 434 GCAGTAGGCTGAGATTATCCATGTTCTTGAATTCGAAGATGAAATATGTCGAAAAGC 493

Db 493 GCGATTAGTGTGGGTGTTCTATGCACTGAGATTGAGAGTGAACATTATTAAGAA 552

Qy 494 CTACATTCAGACGGTGGATGAACCTTGACT---TCCAGMATCTGTAATACAGCCT 550

Db 553 TTGCATTCAGACGGTGAAGAGCTCAACTTATATCTCTGTATTTGACCTATTG 612

Qy 551 TATGAAGTGAACAATATGAAGATTCGCTATGATGAAGAGCTGACGCTAGAGACC 610

Db 613 GAGAACTGAGACAACCTCTGAAAGGCAATTCGATGTAAGGAAATGACATTGCAACC 672

Qy 611 AGTACTACTCTGTTTCTATTTGGAGAAATGCAATGCCACGATTTTTCGCTAATTG 670

Db 673 AAGAGATTCTTATGTTCTTTTGGAGAGTGTGTCACAGAGATTTTGTGCTTATCT 732

Qy 671 AACCAAAATTCGTGTCAGAGCAGCGAGATGACCATTTGATATGTTTCATTACACT 730

Db 733 AATACATCGGTGTCAGAGCAGCGCAATATGATGATTTGAAATGTTTCATTACAGC 792

Qy 731 GATGAATTTGGTAATGCGGATATCTTAGAACCAACCTATCTGTGTTGGAAAGACT 790

Db 793 GATGATTCACAAAGGGGATATCCTGAGGCAACTTATCAGGTGTGCCAAGAGTTA 852

Qy 791 CATGGGACCTGATACAGATTCACGCAATCAGTATCTGTTTCTGAGGTTCTCTGGAAGGC 850

Db 853 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 912

Qy 851 TGAAGATCTGTGCTGTAATCTTAAAGCCGAGGTGTGATGATGATGATGATGATGATGAT 910

Db 913 TGAAGATCTGTGCTGTAATCTTAAAGCCGAGGTGTGATGATGATGATGATGATGATGAT 972

Qy 911 ATTGATTAAGCTTGGGCTGAGAGAAATTAAGGTATGAGAAAGTGTGATGATGATGAT 970

Db 973 ATTGATTAAGCTTGGGCTTGAAGAGATTAAGGTATGAGAAAGTGTGATGATGATGAT 1032

Qy 971 ACTTGATTCACAAATATCTACCCACATGCAAGACGTTTCCATATCTTAACATTTGAAG 1030

Db 1033 ACATGTGACCTACTATATATTAAGAGAGCTCACACAGTACATATCTGACATTTGACGAA 1092

Qy 1031 GCACAGAACTGCTTATTTTGGTGTGCTGAGTTTGGATCGACATGATGAGACCTGCT 1090

Db 1093 GCAAGCCAGCTGATTTATTTTGGTGTGCTGAGTTTGGATCGACATGATGAGACCTGCA 1152

Qy 1091 AGAAGAGGTGATATTCAGTTAGGATTAAAGATTCATCAACCTTAAGCTCCAGGACCC 1150

Db	1153	AAGAGGGGAGATTCCGTTAGGGTTAAAGTTCTTATTAACCTTAAGGCTCTGGAAT	1212
Db	1151	CTTATTACACAAAGAGACATGGAATTAAGGTCTGGTTGTACTAACTTAGCTACTGCTC	1210
Db	1213	AATCACTATAAAACAAGACATACCAAGG-----TATTCTAAAGCAATTTCTTG	1266
Db	1211	AAGTCAATATGCATATGTTGACATTTGAGCACTCGATGCTTGGTCAGTATGTTT	1270
Db	1267	AACGTAATGTGCATCTGTAATATGCAAGACCCGGAATGCTTGGTCAAGTGGCTT	1326
Db	1271	CTGGCAAGGGATCAGGTATTTGCTAATTTGAAGATCTATGTATCTGGATTTGTCT	1330
Db	1327	CTTGGCAAGGTAATTCGATAT-----TTGAGAGGTTGGCAATTCCTGATATTTGTT	1380
Db	1331	GCTACAGTGAAGTAGTGTCTGTGTCACTTATGCATCAAGATCTGAGTAGGGA	1390
Db	1381	GCCACTATGTGAAGTCAGTATATCTCTGACTGATCTTCAAACTTTGGAGCAGADA	1440
Db	1391	CTGATACAGCAGGCAAGTGAATTGTACCATGTAGTTGAAGCTTGAGAAAATACCAAT	1450
Db	1441	CTGATTCAAAG-----GAGCTTGACATGATGTTGATGACAACTGGAGAAAATTCAGTT	1494
Db	1451	GTTTCGTACTTCACACAGAGGGCAGATATTTCACTTAATGGAAATGGAGCAATCGCT	1510
Db	1495	GTGAATCTTTTAAAGGAAGACATCATCTCTTAATTTGGAAATGTTCAACATTTCTCC	1554
Db	1511	CTGATACTGAATAAGCCGAGCGTGTGTGAGGAAAAGTGGGTTATATGTTCAATGATC	1570
Db	1555	CTGATTTTAAAGAGGCGTTTCATGTTCTTATACCAAGAGTGTCAATGTCCAGATGATA	1614
Db	1571	TGGCAAGAGGCTCAAAGTTAATCATGTGCTGATAGTCCATATATAGCATGCAAGGCA	1630
Db	1615	TCCACAGGACATCCAAAGTAAACATTTCTTTAATGTAACCAAGCTGAACGTGAAGT	1674
Db	1631	CTCGTAGAAGCCTTCATCAGCGCTTTTGAAGACATGCTCTATCACA	1680
Db	1675	TGTGTTCAAGGCTTTTCACAATCTCTTTTGAAGGCGGTACCTCTCAGA	1724

RESULT		13
LOCUS	AY089207	
DEFINITION	Arabidopsis thaliana clone 9468 mRNA, complete sequence.	1933 bp MRNA linear PLN_14-APR-2003
ACCESSION	AY089207	
VERSION	AY089207.1	GI:21407981
KEYWORDS	FLI CDNA.	
SOURCE ORGANISM	Arabidopsis thaliana (chale crese)	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
AUTHORS	Spermatocytia; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosoids II; Brassicales; Brassicaceae; Arabidopsits. 1 (bases 1 to 1933) Haas,B.J., Volforsky,N., Town,C.D., Troukhan,M., Alexandrov,N, Feldman,K.V., Flavell,R.B., White,O. and Salzberg,S.L.	
TITLE	Full-length messenger RNA sequences greatly improve genome annotation	
JOURNAL MEDLINE PUBLISHED BY	Genome Biol. 3 (6), RESEARCH0029 (2002) 22088475 12093376	
REFERENCES AUTHORS	2 (bases 1 to 1933) Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.	
TITLE JOURNAL REFERENCE	Full-Length cDNA from Arabidopsis thaliana Unpublished 3 (bases 1 to 1933)	
AUTHORS	Brover,Y., Troukhan.M., Alexandrov.N., lu.Y.-P., Flavell.R. and Feldmann.K.	
TITLE COMMENT	Direct Submission Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of	

This set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or Iaer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genest carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones and sequence assembly.

FEATURES

Source

OFFICE

Query Match

Matches 1007; Conservative 0; Mismatches 401; Indels 22; Gaps 5;

254 ATGGAAGGATTTGGGGATCAGCTCAGCGTGGTGATGAAGTTCCGGGGGCTCCTCGGTTCG 313

280 ACCAGCTACAGCTGATGAGCTTCTGCTG 339

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

100

[illegible]

D6 400 GTCATTGTTCTTCTGGCAAGGAATCTACCAACAATCTCTTGCTTGGGAGAATAA 459

434 GCAGTAGGGTGTGGAGTTATCCATGTTTCTGAATCGAAGAGTGGAAATATGGTCAAAGC 493

Db 460 GCGTTAGTTGTGGTGTTCCTAATGCATCTGAGATTGAGGAGTTGAGCATTATAAGGA 519

494 CTACATATCAAGACGGTGGATGAAGTTGACT--TCCAAGNATCTGTAATACAAAGCCTT 550

520 TTGCATATTCAGACCGTGAAGAGCTCAACATTGATCCCTCTGTATTTTGACCTATTTC 579

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

[illegible]

DB 639 AGAGATTAATTAGTCTCTTTGGAGAGTGTGTCTACAAAGGATTTGGCTTATCTT 050

671 AACAAATTCGTCTCAAGCAGCGCAGTATGACGCATTGTGATTTGGTTTCATTTACACT / 30

Db 699 AATACCATCCGTGTCAAGCACCATAATGATGCATTTGAAATTGTTTCATTACAACG 758

731 GATGAATTTGGTAATGCGGATATCTTAGAAGCAACCTATCCTGCTGTTGCCAAGAGACTT 790

Db 759 GATGATTTCAACAATGGGATATCTTGAGGCACTTATCGAGCTGTTGCCAAGAGATTA 818

791 CATGCGGACCTGCGATTCAAGGATCAAGGATTAACCTGGTTGTTACCTGGGTTCCCTTGGGAAAGGCG 850

[illegible]

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

[illegible]

911 A11GGTAAGCCT1GGAC1GAGAGMM11CAGG1A1GGMMGAG11GH1G1G1M11 5'0

Db. 939 ATTGTAAGCCGTTGGGTTGAAAGAGATTCAGGTGTGGAAGATGTCGATGGTGTCTA 998

QY 971 ACTTGATCCAAATATCTACCAATGCAAGACTGTTCCTACTTAACTTTGAAG 1030
 Db 999 ACATGTGACCTCTATATATTAAGAGCTACACAGTACCAATCTGACATTTGACGAA 1058
 QY 1031 GCCACAGACTTGCTATTTGGTGTCTGAGTTTTCATCCACATCGATGAGACTGCT 1090
 Db 1059 GCAGCCGACTGACTTATTTGGTGTGACAGGCTTGACACCAAGTCATGACACGCA 1118
 QY 1091 AGAGAGGTGATTTTCAGATTGGGTGAAGATTATCAACCTTAAAGCTCCAGGACC 1150
 Db 1119 AGAGAGGTGATTTTCAGATTGGGTGAAGATTATCAACCTTAAAGCTCCAGGACT 1178
 QY 1151 CTATATACCAAGCAAGACATGATTAAGGTGTGTTGTTACTAATAGATAGTCTC 1210
 Db 1179 ATATATCTTAAACCAAGACATGATTAAGGTGTGTTGTTACTAATAGATAGTCTC 1232
 QY 1211 AAGTCAATGTCACTATGTTGACATTTGACACTCGAGTCTTGTGCTATGTTT 1270
 Db 1233 AAAGTAAATGTCACTATGTTGACATTTGACACTCGAGTCTTGTGCTATGTTT 1292
 QY 1271 CTGGCAAGGTATCAGGTATTTGCTATATTTGAAGATCATATCTGTGATTTGTT 1330
 Db 1293 CTGGCAAGGTATTTGCTATATTTGAAGATCATATCTGTGATTTGTT 1346
 QY 1331 GCTACCAAGTAAATGTTGTTCTGTGCTATGTTGATTCATCAAGATCTGAGTGAAGAA 1390
 Db 1347 GCCACTAGTAAAGTCAATGTTCTGTGACATGATTCCTTCAAACTTTGAGCAGAGAA 1406
 QY 1391 CTGATACAGCAGCAAGTAACTTACCATGTTAGTTGAAGCTTGAAGAAATAGCAATT 1450
 Db 1407 CTGATTCACACAG-----GAGCTGATCTATGTTGAGAACTGGAAGAAATAGCAATT 1460
 QY 1451 GTTCGTCTACTCAGCAGAGGCGATTAATTCATCTATTCGAAATGTTGAGCAATCTCT 1510
 Db 1461 GTGAACTCTCTAAAGAGAGAGCAATCATCTCTTAATTTGGAAATGTTCAACATCTCC 1520
 QY 1511 CTGATCTAGAAAGACGGGACGTGTCTGAGAGAAAGTGGGTTAATGTTCAATGATC 1570
 Db 1521 CTGATTTTGAAGAGGCGCTTTCATGTTCTTTATTAACAAAGGTGTAAGTCAATGATA 1580
 QY 1571 TCGCAGAGAGCGTCAAGGTAAATGCTGATGATGATGATGATGATGATGATGATGATG 1630
 Db 1581 TCGCAGAGAGCGTCAAGGTAAATGCTGATGATGATGATGATGATGATGATGATGATG 1640
 QY 1631 CTGTAAGAGCCCTTATCAGGCTTCTTGAAGCATGTCCTATCA 1680
 Db 1641 TGTGTCAGGCTCTTCAAAATCTTCTTGAAGCGGTGACCTCTAGA 1690

RESULT 14
 ATU62020 1923 bp mRNA linear PLN 15-JUL-1997
 LOCUS Arabidopsis thaliana lysine-sensitive aspartate kinase mRNA.
 DEFINITION complete cds.
 ACCESSION U62020
 VERSION U62020.1 GI:2257742
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1923)
 Tang, G., Zhu-Shimoni, J.X., Amir, R., Zchori, I.B. and Galili, G.
 Cloning and expression of an Arabidopsis thaliana cDNA encoding a
 monofunctional aspartate kinase homologous to the lysine-sensitive
 enzyme of Escherichia coli
 JOURNAL Plant Mol. Biol. 34 (2), 287-293 (1997)
 MEDLINE 97351562
 PUBMED 9207844
 REFERENCE 2 (bases 1 to 1923)
 TANG, G., ZHU, X.-Z. and GALILI, G.

TITLE Direct Submission
 JOURNAL Submitted (25-JUN-1996) Plant Genetics, The Weizmann Institute of
 Science, Rehovot 76100, Israel
 FEATURES
 source Location/Qualifiers
 1..1923
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 117..1751
 CDS
 /note="similar to E. coli lysine-sensitive aspartate kinase"
 111, Swiss-Prot Accession Number P08660
 /protein_id="AAB63104.1"
 /db_xref="gi:2257743"
 /translation="MASIOLYGVTPGLATLSKRLERFASKGACFSVTLPSAVFRDV
 /BSCNIGLRVSCALRVLDLORKEPTCDSSGKELTCVMKFGSSVGSABRMKEV
 ANLIFSPEDRPPVYL SAMGKTINKLKAGEKATVGTWVGSLEBLSFKELRLTA
 HBLGVETVIERKHLEGLHOLKGISMKELTLTRDYLVSFGCKSTRFLPSALYKIG
 HKARQYDAFEIGFTTDDFTNADILEATYPAVSKTLVDMSKENAVPVATYIGKMR
 SCATITLFGGSDLTATTIGALGLEIYQVKVDVLTDPNIYEQASVPYLTDE
 AAEIAYFGAQLHPLSRPADGDIPIRVKNSYNPFLPFTVITRSBMSKAVLSIVL
 KRNVTMLDIASRLMGYFLAKVFTFEDLGSIVDVATSEVSLITLDPALMRE
 LIORNEIDNIVERLEIAVAKLIORSIIISLIGVKSLSLEKVFQVRSNVMVQ
 MISQASRKNISLIVNDEBAQCVRALHSAFFEND"
 ORIGIN
 Query Match 34.2%; Score 678.4; DB 8; Length 1923;
 Best Local Similarity 69.1%; Pred. No. 6.4e-154;
 Matches 976; Conservative 0; Mismatches 422; Indels 15; Gaps 3;
 QY 260 GGATTGGGGGATCAGCTCAGCGTGTGATGAAGTTCCGGGGGCTCGTGTGTGGGCC 319
 Db 345 GGAAGTGGCAAGATTAATCAATGTAAATTTGGCGGTTATCAGTGGATACGCA 404
 QY 320 GCGAGATGGCTGAGAGTGGCGGCTCATCTCTGACGTTCCCGAGAGCGCCCGTGTG 379
 Db 405 GAGAGATGMAAGAGTTGCCAATCTTATATTCAGCTTCCAGATGAGAGCCCGCTATT 464
 QY 380 GTTCTCTGCCATGGGAAAAACCAACAACCTTCTCTGTTGAGAGAGGACGTA 439
 Db 465 GTGCTATAGCATGGGAAAGACATTAATAGCTTTTGAAGCTGAGAGAGGCTGTT 524
 QY 440 GGGTGGAGGTATCATGTTTCTGAATCCAGAGTGAATATGTCACAAAGCGCTCAT 499
 Db 525 ACTTGGCGGTGACATGAGATGAGATTTGAAGATTTGAGCTTTATTAAGAACTGCAT 584
 QY 500 ATCAAGCGGTGATGAACCTTGACT--TCAGAAATCTGTAATACAGCTTTATGAA 556
 Db 585 TTAAGAACTGTCATGAGCTTGGAGTGAACAACGCTTATTAAGAAACCTGAAAGA 644
 QY 557 CTGGAACAATTAAGAAAGTATGCTATGATGAAGAGCTGAGCGCTAGAGCAAGTAC 616
 Db 645 CTACACCAAGCTTCTAAGAGCATATCAATGATGAAGAGCTAACCTTACGCAAGAGAC 704
 QY 617 TACCTGTTTCAATTGGAGATGATGTCACAGATTTTTCGTGTTATTGAACAAA 676
 Db 705 TACTTGTTTCAATTGGAGATGATGTCACAGAGCTTTTCTGCGTATCTCAATAAA 764
 QY 677 ATTGCTGTCAGGACCGGAGATGACCATTTGATATTTGTTTCAATTAACAATGAGAA 736
 Db 765 ATTGCGCAAAAGCCGTCATATATGATGATTTGATGATGATTTATCACACAGAGAC 824
 QY 737 TTGGAATGCGGATATCTTGAAGCAACCTATCTGTTGTTGGAAGAGCTTCATGGG 796
 Db 825 TTACTAATGCGATATCTTGAAGCAACATACCTGCGCTCAAAAACATGTTGGGC 884
 QY 797 GACTGATACAGATTCACGAGTACCTGTTTCTGCTGCTCTTGGAGAGGCGCTGAAA 856
 Db 885 GACTGAGCAAGAAATGACAGTCCCTGTTTCTGCTGCTACCTGGAAGAGGATGAGG 944
 QY 857 TCTGTGCTGTAATCTTATGAGCGAGGTGTGATGATGATGATGATGATGATGATGATG 916

Db 945 TCTTGCTATCTACGCTGGGTGGGAGTATTTGACGCTACACGATTGGG 1004
 QY 917 AAAAGCTTGGAGCTAGAGAAATTCAGGTATGAAAAGTGTGTACTTACTTGT 976
 Db 1005 AAAAGCTAGAGCTGGGAAATTCAGGTATGAAAAGTGTGTACTTACTTGT 1064
 QY 977 GATCCAAATATCTACCCCATGCAAAAGCTGTTCATCTTAATTTGAAGGCCACA 1036
 Db 1065 GATCCAAATATCTACCCCATGCAAAAGCTGTTCATCTTAATTTGAAGGCCACA 1124
 QY 1037 GAATCTGCTTATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1096
 Db 1125 GAGCTTGTCTTATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1184
 QY 1097 GGTGATATTCAGTTAGGTTTAAATTCATCAACCTTAAAGCTCCAGGACCTTAT 1156
 Db 1185 GGGCAGATCTGTGAGGTTTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1244
 QY 1157 ACCAGCAAAAGACATGATPANGTCTGGTGTGTGTGTGTGTGTGTGTGTGTGT 1216
 Db 1245 ACCAGTCAAGACATGATPANGTCTGGTGTGTGTGTGTGTGTGTGTGTGTGT 1298
 QY 1217 AATGTCACTATGTGTGACATTTGTGACATCTGGATGCTTGTGTGTGTGTGTGT 1276
 Db 1299 AATGTCACTATGTGTGACATTTGTGACATCTGGATGCTTGTGTGTGTGTGTGT 1358
 QY 1277 AAGGTATCAAGTATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1336
 Db 1359 AAGGTATCAAGTATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1412
 QY 1337 AGTGAAGTATGT 1396
 Db 1413 AGTGAAGTATGT 1472
 QY 1397 CAGCAGGCAAGTGAATTCATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1456
 Db 1473 CAGCAGGCAAGTGAATTCATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1532
 QY 1457 CTACTTCAGCAGAGGCGCATTAATTTCACTTATCGAAAATGTGTGTGTGTGTGT 1516
 Db 1533 CTACTTCAGCAGAGGCGCATTAATTTCACTTATCGAAAATGTGTGTGTGTGTGT 1592
 QY 1517 CTGAAAAGACGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1576
 Db 1593 TTAGAGAGGTGTTCAGATATTTTGAAGCAATGGAATGTGTGTGTGTGTGTGT 1652
 QY 1577 GAGAGCTCAAGGTTTAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1636
 Db 1653 GAGAGCTCAAGGTTTAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1712
 QY 1637 GAAAGCCCTTCATGAGCGTTCTTTTGAAGCAAT 1669
 Db 1713 AGGAGCTTCACCTCCGCTTTTGTGACCGAT 1745

RESULT 15
 LOCUS AX180354 1658 bp DNA linear PAT 06-AUG-2001
 DEFINITION Sequence 15 from Patent WO0146393.
 ACCESSION AX180354
 VERSION AX180354.1 GI:15132296
 KEYWORDS
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticaceae; Triticum.
 1
 Pato,S.C., Famodu,O.O. and Thorpe,C.J.
 TITLE Apatate kinase
 JOURNAL Patent: WO 0146393-A 15 28-JUN-2001;
 E.I. DU PONT DE NEMOURS AND COMPANY (US)
 FEATURES
 Location/Qualifiers

source 1..1658
 /organism="Triticum aestivum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4565"
 ORIGIN
 Query Match 34.0%; Score 674.6; DB 6; Length 1658;
 Best Local Similarity 72.8%; Pred. No. 5,3e-153;
 Matches 930; Conservative 0; Mismatches 326; Indels 21; Gaps 4;
 425 GAGAGAAAGGAGAGTGGGTGTGAGTATTCATCTTTCTGAAAATCCAAAGTGAATANG 484
 Db 4 GGCAGAGAGGAGGAGAGTGGGTGTGAGTATTCATCTTTCTGAAAATCCAAAGTGAATANG 63
 QY 485 GTCAAAAGCCATCATCAAGACGGTGGATGAATCTTGACTTCAAGNATCTGTATAGA 544
 Db 64 ATCAAGAGCTTCATCTCAAGACCATGATGAGCTTGCCATGATGCTTCATTTTCA 123
 QY 545 AGCTTTT---ATGACTGTGAGCAATTTGAAAGTATGCTATGATGAAAGAGCTGACG 601
 Db 124 GGTTTTGTGAGAGTGTGAGCAATGCTCAAGGAGTGTGTATGATGAAAGAGCTGACT 183
 QY 602 CTTAGGACCATGACTTACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 661
 Db 184 CTTAGGACCATGACTTACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 243
 QY 662 GCTTATTTGAACAAATTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 721
 Db 244 GCATTTTGAATTAACATGAGGAAAGAGCAAGACATGATGATGCTTTTGTATCTTGCT 303
 QY 722 ATTCACTGTATGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 781
 Db 304 ATTAACACTGTGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 363
 QY 782 AAGAGACTTCATGGGAGCTGTGATCAGAGATCCAGCGATACCTGTGTGTGTGTGT 841
 Db 364 AAGAGCTTCATGGGAGCTGTGATCAGAGATCCAGCGATACCTGTGTGTGTGTGT 423
 QY 842 GGAAGGCTGGAATCTGT 901
 Db 424 GGAAGGCTGGAATCTGT 483
 QY 902 GCTAACACATTTGTGAAAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 961
 Db 484 GCTAACACATTTGTGAAAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 543
 QY 962 GGTGTACTTACTTGT 1021
 Db 544 GGTGTACTTACTTGT 603
 QY 1022 TTTGAAGGCGCACAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1081
 Db 604 TTTGAAGGCGCACAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 663
 QY 1082 AGACCTGTGAGAGAGTGTATTTCCAGTTAGGGTTAGATTCATCAACCTTAAGCT 1141
 Db 664 GACACAGCAGAGAGAGTGTATTTCCAGTTAGGGTTAGATTCATCAACCTTAAGCT 723
 QY 1142 CCAGGACCTTATTTATCCAGCAAGAGCATGTATANGTCTGTGTGTGTGTGTGTGT 1201
 Db 724 CCGGACCTGTGTACTTAAACAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 777
 QY 1202 ATAGTGTCAAGTCAATATGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1261
 Db 778 ATTGTCTGAATCAATCAATTTACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 837
 QY 1262 TATGT 1321
 Db 838 TATGT 891
 QY 1322 GATTTGT 1381
 Db 892 GATTTGT 951


```
QY 1382 AGTAGGGAAGTATACAGCAGGCAAGTGAATTGACATGATGTTGAAGAGCTTGAGAA 1441
DB 952 AGTCGTAATTGATCCAGCAG-----GAGCTTGATCATGTAAGTTGAAGAGCTTGAAAG 1005
QY 1442 ATAGCAATTGTTGTTCTACTTTCAGCAGAGGGGATTAATTCACTTATCGAAATGTGAG 1501
DB 1006 ATGCGGTTGTTCACTCTTACAGCAGATCATCATTTCCCTGATAGGGAATGTGCAG 1065
QY 1502 CATGCTCTGATTAAGAAAGAGGAGCTGCTGAGGAAAGTGGGTTAATGTT 1561
DB 1066 AGATCGTCTGATTTTGAGAAAGGCGTTCAATGTTTACGAGAAATGTTTAATGTT 1125
QY 1562 CAGATGATCTCGAAGAGCGTCAAGGTTAACATGTCGTGATAGTCATGATAGCGAT 1621
DB 1126 CAGATGATTTGGAAGGCGGTCAAGGTGAACATCTCCTTGTTGATGACAGCGAG 1185
QY 1622 GCAAAGGCACTCGTAAGAGCCCTTCACTAGGCGTTCTTGAAGAGATGTCCTATCAAA 1681
DB 1186 GCGAAGCAAGTGCCTGCAAGCCCTCCACTCGGCATTCTTGAAGAGGTTTCTTGTCAGAA 1245
QY 1682 GTGGAAGCGAGAGACT 1698
DB 1246 GTAGAGGAAGCGAGCT 1262
```

Search completed: March 23, 2004, 09:36:41
Job time : 7890 secs

CC product procedures. Nucleic acid encoding aspartate kinase is useful
CC for producing a transgenic plant, by transforming a plant cell with a
CC chimeric gene comprising nucleic acid encoding aspartate kinase operably
CC linked to a regulatory sequence and regenerating a plant from the
CC transformed plant cell. The present sequence is cholic.pk002.k6 (F1S) cDNA
CC clone encoding Zea mays aspartate kinase

Sequence 1985 BP; 506 A; 419 C; 535 G; 523 T; 0 U; 2 Other;

Query Match	99.98%	Score 1983;	DB 5;	Length 1985;
100.0%	Prod No. 01			

Matches 1985; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY	1	GCACGAGACAGAGCAGAGAGACTCAGAAATGGCAATCCCACTGGGATCGGCTGCGCGCC	60
Db	1	GCACGAGACAGAGCAGAGAGACTCAGAAATGGCAATCCCACTGGGATCGGCTGCGCGCC	60
QY	61	CGCGCGCCTGCTCCGTCGATACCTTCGCGCGAGCTCTGACAAATTCAGAGATCGGGGTG	120
Db	61	CGCGCGCCTGCTCCGTCGATACCTTCGCGCGAGCTCTGACAAATTCAGAGATCGGGGTG	120
QY	121	CTTCGGTACCCGAAACCGGAGCTCGCGGATTCAGAGAGGTGTCAATGGTGTGCGCGATC	180
Db	121	CTTCGGTACCCGAAACCGGAGCTCGCGGATTCAGAGAGGTGTCAATGGTGTGCGCGATC	180
QY	181	CACCAAGCGTGGGGCCCAAGCAGACGCGGGGAGCGGCGCTCTTGGGGCGCTGTCT	240
Db	181	CACCAAGCGTGGGGCCCAAGCAGACGCGGGGAGCGGCGCTCTTGGGGCGCTGTCT	240
QY	241	CGAGAGGCTCGGGATGAGAGATTGGGGATCAGCTCAGCTGTGTGATGAGATTGCGAGG	300
Db	241	CGAGAGGCTCGGGATGAGAGATTGGGGATCAGCTCAGCTGTGTGATGAGATTGCGAGG	300
QY	301	GTCCTGGGTGTGTCGCGCGCGAGATGGCTGAGGTGGCGGCTCATCTGAAGTTC	360
Db	301	GTCCTGGGTGTGTCGCGCGCGAGATGGCTGAGGTGGCGGCTCATCTGAAGTTC	360
QY	361	CGAGAGCGCCCGGTGTCCTCTCTGCGCATGGGAAAAACAACAACCTTCTCT	420
Db	361	CGAGAGCGCCCGGTGTCCTCTCTGCGCATGGGAAAAACAACAACCTTCTCT	420
QY	421	TGCTGAGAGAAAGGACGTAGGGGTGTGAGATTTCATATTCTGAAATCGAAGGTGAA	480
Db	421	TGCTGAGAGAAAGGACGTAGGGGTGTGAGATTTCATATTCTGAAATCGAAGGTGAA	480
QY	481	TATGTCAAAAGCTTACATATCAAGCGGTGATGAATCTGACCTTCCAAGATCTGAA	540
Db	481	TATGTCAAAAGCTTACATATCAAGCGGTGATGAATCTGACCTTCCAAGATCTGAA	540
QY	541	TACAACCTTTTGAACGTGAGCAACTATGAAAGTATCGCTATGATGAAAGGCTGAC	600
Db	541	TACAACCTTTTGAACGTGAGCAACTATGAAAGTATCGCTATGATGAAAGGCTGAC	600
QY	601	GCTTAGACCAAGTACTACTTGTTCATTGGAGATGACATGTCCACAGATTTTTC	660
Db	601	GCTTAGACCAAGTACTACTTGTTCATTGGAGATGACATGTCCACAGATTTTTC	660
QY	661	TGCTTATTGGAACAAAATTCGTGTCAAGGACGGAGATGACGATTTGATTTG	720
Db	661	TGCTTATTGGAACAAAATTCGTGTCAAGGACGGAGATGACGATTTGATTTG	720
QY	721	CATTACAAGTGAATTTGGTAAATGCGGATCTTAAAGCAACTATCTGCTGTTC	780
Db	721	CATTACAAGTGAATTTGGTAAATGCGGATCTTAAAGCAACTATCTGCTGTTC	780
QY	781	GAGAGACTTCATGGGAGCTGATACAGGATCAGACGATACCTGTGTATTACGGTCT	840
Db	781	GAGAGACTTCATGGGAGCTGATACAGGATCAGACGATACCTGTGTATTACGGTCT	840
QY	841	TGGGAAGGCTGGAATTCGTGTCTGTAACTTCTTAAAGGAGGTGTATGACTGAC	900
Db	841	TGGGAAGGCTGGAATTCGTGTCTGTAACTTCTTAAAGGAGGTGTATGACTGAC	900

[illegible]

Db	125	TTTTCGCGAAGGGATTATTGTAATAT-----TTGAAGATCTATGTAATATCTGTGGATG	1307
Qy	1327	TGTTGGTACACAGGAAGTTAGTGGTTTCTGTGTCTCATGATCCATCCAAAGTCTGGAGTAG	1386
Db	1308	TGTTGTACCAAGGAAGTATGATGTTTCTGTGTCTCATGATCCATCCAAAGTCTGGAGTAG	1367
Qy	1387	GGAACGTATPACAGACGACGCAATGTAACTTGACATGTAGTTGAAGCTTGAGAAAAATAGC	1446
Db	1368	GGAACGTATPACAGACGAG-----GAACTTGACCAATGTAGTTGAAGAGCTTGAGAAAAATAGC	1421
Qy	1447	AATTGTTGCTACTACTTCACGACAGGGGGATTAATTTCACTTATCGGAAATGTGGAGCATTC	1506
Db	1422	AATTGTTGCTCTACTTCACGACAGGGGCATTAATTTCACTTATCGGAAATGTGGAGCATTC	1481
Qy	1507	GTCCTGATATCTAGAAAAGACGGAGCGTGTGCTGATGAGAAAAAGTGGGATTATGTTCAAT	1566
Db	1482	GTCCTCATACTATGAAAAGACGGAGCGTGTGCTGATGAGAAAAAGTGGGATTATGTTCAAT	1541
Qy	1567	GATCTGGCAAGAGGCGTCAAAAGTTAATCATGTGCTGATGATGTCATGATPAGGATGCAAA	1626
Db	1542	GATCTGGCAAGAGGCGTCAAAAGTTAATCATGTGCTGATGATGTCATGATPAGGATGCAAA	1601
Qy	1627	GGCACCTGATAGAAAGCCCTTCATCAGGCGTCTCTTTGAAGCAATGTGCTATCACAATCGA	1686
Db	1602	GGCACCTGATAGAAAGCCCTTCATCAGGCGTCTCTTTGAAGCAATGTGCTATCACAATCGA	1661
Qy	1687	AGCGGAGAACCTACTCTGTGGGCTGATTCAGATGAGCTTCTGTGGGTCAGAGCGTGTATTC	1746
Db	1662	AGCGGAGAACCTACTCTGTGGGCTGATTCAGATGAGCTTCTGTGGGTCAGAGCGTGTATTC	1721
Qy	1747	TGTTATPAGATTCCCACTGCGCTCCATGAAAGCAATGAGCAATTTGATCATTTGATCATGTT	1806
Db	1722	TGTTATPAGATTCCCACTGCGCTCCATGAAAGCAATGAGCAATTTGATCATTTGATCATGTT	1781
Qy	1807	TGCTTGAACCAAGTATGCTTCCAGGTTCTCGAGCCATGACTGCAAAACCTGTGTTCTGT	1866
Db	1782	TGCTTGAACCAAGTATGCTTCCAGGTTCTCGAGCCATGACTGCAAAACCTGTGTTCTGT	1841
Qy	1867	TTTAAAGACTGTTTGGACACCAAGTACGCTCGAGCAACCGATTTGTCAACAGATGGCAAG	1922
Db	1842	TTTAAAGACTGTTTGGACACCAAGTACGCTCGAGCAACCGATTTGTCAACAGATGGCAAG	1903
Qy	1927	CCGTGTATTAATTAATTCACACTGCTCTATTCATATTAATTAATTAATTAATTAATTC	1978
Db	1902	CCGTGTATTAATTAATTCACACTGCTCTATTCATATTAATTAATTAATTAATTAATTC	1953
RESULT 3			
ABZ13939			
ID	ABZ13939	standard; DNA; 1710 BP.	
XX	ABZ13939;		
AC			
XX	21-JAN-2003	(first entry)	
DT			
XX			
DE	Arabidopsis thaliana stress regulated gene SEQ ID NO 1744.		
XX			
KW	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.		
XX			
CS	Arabidopsis thaliana.		
XX			
PN	MO200216655-A2.		
XX			
PD	28-FEB-2002.		
XX			
PF	24-AUG-2001; 2001MO-US026685.		
XX			
PR	24-AUG-2000; 2000US-0227866P.		
ER	26-JAN-2001; 2001US-0264647P.		
XX	22-JUN-2001; 2001US-0300111P.		
PA	(SCRT) SCRTIPS RES INST		
PA	(SYGN) SYNGENETA PARTICIPATIONS AG.		

Pt	XX	Harper JF, Kieps J, Wang X, Zhu T;
Xx	XX	WPI; 2002-304127/34.
Xx	PT	Identifying a stress condition to which a plant cell has been exposed and
Pr	XX	producing plants with increased tolerance to these abiotic stresses.
Ps	XX	Claim 144; SEQ ID NO 1744; 577bp + Sequence Listing; English.
Cc	XX	The invention relates to identifying a stress condition to which a plant
Cc	CC	cell has been exposed, comprising: (a) contacting nucleic acid
Cc	CC	representative of expressed polynucleotides in the plant cell with an
Cc	CC	array or probes representative of the plant cell genome; and (b)
Cc	CC	detecting a profile of expressed polynucleotides in the plant cell
Cc	CC	characteristic of a stress response. The method is useful in the
Cc	CC	production of transgenic plants, cells and seeds and in producing plants
Cc	CC	with increased tolerance to abiotic stress. The present sequence is that
Cc	CC	of an Arabidopsis thaliana stress regulated gene (AB21296-AB217574) used
Cc	CC	in methods of the invention. Note: the sequence data for this patent is
Cc	CC	not represented in the printed specification but is based on sequence
Cc	CC	information supplied to Derwent by the European Patent Office
Sq	XX	Sequence 1710 BP; 478 A; 323 C; 426 G; 483 T; 0 U; 0 Other;
Qy	Query Match	35.6%; Score 707.6; DB: 6; Length 1110;
Db	Best Local Similarity	70.5%; Pred. No. 1.5e-198;
Db	Matches 1008; Conservative	0; Mismatches 401; Indels 21; Gaps 4
Oy	254	ATGAGAGATTGGGGGTACGCTCAGCGTGTGAAGAATTGGGGGGTCTCGTGTCG 313
Db	235	ACGAGATGATGAGAGAGGTATCACTGCGTGAAGATTGGTAGATCTTCGGTGGCG 294
Oy	314	TGGGCCCGAGAGATGCTGAGGTGGCCGGCTCATCTGACGTTCCCCGAGAGCGCCC 373
Db	295	TGAGCTGAGAGATGAGAAATTGCTAATTGATTGATTTTCCGAGAAAGATCCC 354
Oy	374	GTCGTGTTCTCTGCSCATGGGGAACAACCAACAACCTTCTCCTGTGAGAGAG 433
Db	355	GTATGTTCTCTCTGCTATGGGAAAATCAACAATCTCTGCTTGGGGAGAG 414
Oy	434	GCAATGGGTGAGACTTATCATGTTTTCTGAATCGAAGGTGGAATATGTCATAAGC 493
Db	415	GCGGTAGTGGTGGTTTCTATGCACTTGAATGAGATGAGGTGACATTATTAAGAA 474
Oy	494	CTCATATCAGACGGTGTGTAATTGAGACT--TCCAAGNATCTGTAATACAGCCT 550
Db	475	TTCATATTCAGACGGTGAAGCTCAACATGATCCCTGTATTATTTGACCTATTG 534
Oy	551	TATGAATCGAGCAACTATTGAAGGATTCGCTATGATGAAGAAGCTGACGCTAGACC 610
Db	535	GAGAACTGGAGCACTCCCTGAAGAAGCATTCGATGAGAGGAATTGACACTCGAAC 594
Oy	611	AAGCATACCTGTTTCAATTTGGAAGAAGCATGTCACAGATTTTTCGCTATTG 670
Db	595	AGGATTACTTAATCTCTTTTGAAGAGTGTGTCTACAGAGATTTTTCGCTACTT 658
Oy	671	AACAAATTCGTGTCAGGACAGGACGAGTAGTAGCATTGATTTGTTTCATTACACT 730
Db	655	AATACATCGGTCTCAAGCAGCCAAATATGATGATTGAATTTGTTTCATTACACG 714
Oy	731	GATGAATTGTAATCCGGAATCTTATAAGCAACTATCCTGCTGTGGGAAGACTT 790
Db	715	GATGATTTACAAATGGGATATCTCGAGGCACTTATCCAGCTGTTCGCAAGATTA 774
Oy	791	CATGGGACTGATACAGATCACACGATACCTGTTGTTACTGGGTTCTTTGGAGGCG 850
Db	775	TATGATGATGATGATGATGATCTCTGCTCTTCTATATGAACAGGTTTCTTGGAGAGGT 834
Oy	851	TGGAATCTGGTCTCTAATCTATTAGCCGAGGTGTATGACTTGACGCTACACACC 910
Db	835	TGGAATCTGGTCTCTAATCTATTAGCCGAGGTGTATGACTTGACGCTACACACA 894

QY 911 ATTGTAAGCCTTGGACCTGAGAGAAATTCAGTATGAGAAAGTGTGATGCTGACT 970
DB 895 ATTGTAAGCCTTGGACCTGAGAGAAATTCAGTATGAGAAAGTGTGATGCTGACT 954
QY 971 ACTTGTGATCCAAATATCTACCCATGCAAGACTGTTCCATCTTAACATTGAGAG 1030
DB 955 ACATGTGACCTTACTATATTAAGGACCTACACGATACCATCTGACATTTGAGAA 1014
QY 1031 GCCACAGAACTTCTATTTTGTGCTGAGTTTGTTCATCCACATGATGAGACTGCT 1090
DB 1015 GCAGCGGACCTAGCTTATTTTGTGCAACGGCTTGTGCAACCATGATGAGCCAGCA 1074
QY 1091 AAGAGAGGTGATATTCAGTTAGGTTAAGATTATACACCTTAAAGCTCCAGGACC 1150
DB 1075 AAGAGAGGTGATATTCAGTTAGGTTAAGATTATACACCTTAAAGCTCCAGGACT 1134
QY 1151 CTTATACCAAGCAAAAGACATGATTAAGGTCTGTTGTTACTACTAGCATAGTCTC 1210
DB 1135 ATCATCTCTAAACAAAGACATGACCAAGS-----TATTTAGAGCATTTGTTCTG 1188
QY 1211 AAGTCAATGCACTATGTTGACATTTGTGACACTGAGATGCTTGTGATGCTTTT 1270
DB 1189 AAACGTATGATACCATGCTGATATGACAGACCCCAATGCTTGTGCAAGTTGGCTTT 1248
QY 1271 CTGGCAAGGTATCAGTATTTGCTATATGTAAGATCTATGATATCTGTGATTTGTT 1330
DB 1249 CTTGCAAGGTATTTCTCGATAT-----TTAGAGCTTGTGCAATTTCCGTATGTTGTT 1302
QY 1331 GCTACCAAGTAAATGATTTCTGTGCTACTGATCCATGCAAAATCTGAGTAGGAA 1390
DB 1303 GCACACTAGGAAGTCAATATCTCTGACACTGATCTCTTCAAACTTGGAGAGAGAA 1362
QY 1391 CTGATACAGGAGCAAGTAACTGACCACTGATGTAAGAGCTTGAAGAAATGCAAT 1450
DB 1363 CTGATCAACAG-----GAGCTGATCTGATGTTGAAGACTGGAAGAAATGCAAT 1416
QY 1451 GTTCGTCTACTTACAGAGAGGCGATTAATTTCACTTATCGAAATGTGAGCAATGCT 1510
DB 1417 GTGATCTTAAAGAGAGCAATCATCTCTTAATTTGGGAATGTTCAACATTTCTCC 1476
QY 1511 CTGATCTGAAAGACGGGAGCTGTGCTGAGAGAAAGTGGGTTAAGTTCAGATGATC 1570
DB 1477 CTGATTTAGAGAGGCGCTTATGTTCTTTATACCAAGGTAAATGTCAGATGATA 1536
QY 1571 TCGCAGAGAGCGTCAAGGTTAACATGTCGCTGATAGTCCATGATGAGTCAAGGCA 1630
DB 1537 TCGCAGAGAGCATCCAGGTAAACATTTCTTTATAGTAAAGCAAGCTGAAAGCTGAGGT 1596
QY 1631 CTGCTAGAGCCCTTATCAGGCGTTCTTTAAGACGATGTCCTATCA 1680
DB 1597 TGTGTTCAGGCTCTTCAAAATCTTTCTTGAGAGCGGTGACCTCTGAGA 1646

RESULT 4

AAC37959 standard; DNA; 1933 BP.

AAC37959;

17-OCT-2000 (first entry)

Arabidopsis thaliana DNA fragment SEQ ID NO: 19281.

Hybridisation assay; genetic mapping; gene expression control;
protein identification; signal transduction pathway; metabolic pathway;
promoter; termination sequence; ss.

Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130448P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134321P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134376P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.

	Query Match	35.0%	Score 695.6	DB 3	Length 1933
	Best. Local Similarity 70.4%		Pred. No. 5,7e-195		
	Matches 1007	Conservative 0	Mismatches 401	Indels 22	Gaps
QY	254 ATGAGAGGATTGGGGATCAGCTCAGCGTGGTATGAAGTTGGGGGGTCTTGGTGTG	313			
Db	280 ACGAGAGTATGATGAAGAGGTATCAGCTGCGTATGAAGTTTGTGGATCTTGGTGGCG	339			
QY	314 TCGGCCGAGAGATGCGCAGAGTGGCGGCGCTCATCTCGACGTTCCCGAGGAGCGCCCC	373			
Db	340 TCAAGCGAGAGATGAAGAAAGTTGCTGAATTTGATTTTGACTTTTCGGAAGAAATGCC	399			
QY	374 GTCTGTGTTCTCTCGCCANTGGGAAAAACCAACAACCTTCTCCTTGTGAGAGAAG	433			
Db	400 GTCATTTGTTCTCTGTGATGGGAAAAACCAACAATCTCTGCTTGGGAGAGAG	459			
QY	434 GCAGTAGGGTGTGAGATTATCCATGTTTTCGAATTCGAAAGAGGAATTTGTCAAAAGC	493			
Db	460 GCGGTATGTTGGTGTGTTCTTAATGCACTGAGATTGAGAGTTTGAACATTAATAGGA	519			
QY	494 CTACATATCAAGCGTGGATGAACCTTGACT--TCCAGAACTCTGATATCAAGCCTT	550			
Db	520 TTGCATATCAGAACGCTGAAGAGCTCAACTGTATCCCTGTATATTTTGAACCTATTG	579			
QY	551 TATGAATCGAGACACTTGAAGATGATCGTATGATGAAGAGCGTACGCTAGAC	610			
Db	580 GAGGAATGAGCAACTCTTAAGAGCAATTCACATGAT--AAGAAATTGACTTGAAC	638			
QY	611 AGTACCTACCTGTTTCAATTTGAGAAATGATGTCACAGATTTTTCGTATTG	670			
Db	639 AGAGTATTACTTGTCTTTTGGAGAGTGTGCTCAAGAGATTTTGTGCTTATCTT	698			
QY	671 AACAAATTCGTGTAAAGCAGGAGATGAGCAATTTGATTTTCAATTAACCT	730			
Db	699 AATACATGGGTGTAAGGACGCCAATATGATGCATTTTGAATTTGGTTTCATTACAG	758			


```
PR 22-JUN-1999; 99US-0139899P;
PR 23-JUN-1999; 99US-0140353P;
PR 24-JUN-1999; 99US-0140354P;
PR 24-JUN-1999; 99US-0140695P;
PR 28-JUN-1999; 99US-0140823P;
PR 29-JUN-1999; 99US-0140991P;
PR 30-JUN-1999; 99US-0141287P;
PR 01-JUL-1999; 99US-0141842P;
PR 01-JUL-1999; 99US-0142154P;
PR 02-JUL-1999; 99US-0142055P;
PR 06-JUL-1999; 99US-0143390P;
PR 08-JUL-1999; 99US-0144803P;
PR 09-JUL-1999; 99US-0144920P;
PR 12-JUL-1999; 99US-0144972P;
PR 13-JUL-1999; 99US-0144974P;
PR 14-JUL-1999; 99US-0144975P;
PR 15-JUL-1999; 99US-0144005P;
PR 16-JUL-1999; 99US-0144085P;
PR 16-JUL-1999; 99US-0144086P;
PR 19-JUL-1999; 99US-0144325P;
PR 19-JUL-1999; 99US-0144331P;
PR 19-JUL-1999; 99US-0144332P;
PR 19-JUL-1999; 99US-0144333P;
PR 19-JUL-1999; 99US-0144334P;
PR 19-JUL-1999; 99US-0144335P;
PR 20-JUL-1999; 99US-0144352P;
PR 20-JUL-1999; 99US-0144632P;
PR 20-JUL-1999; 99US-0144884P;
PR 21-JUL-1999; 99US-0144814P;
PR 21-JUL-1999; 99US-0145086P;
PR 21-JUL-1999; 99US-0145087P;
PR 21-JUL-1999; 99US-0145088P;
PR 22-JUL-1999; 99US-0145089P;
PR 22-JUL-1999; 99US-0145089P;
PR 22-JUL-1999; 99US-0145089P;
PR 22-JUL-1999; 99US-0145192P;
PR 23-JUL-1999; 99US-0145145P;
PR 23-JUL-1999; 99US-0145214P;
PR 23-JUL-1999; 99US-0145224P;
PR 26-JUL-1999; 99US-0145276P;
PR 27-JUL-1999; 99US-0145913P;
PR 27-JUL-1999; 99US-0145918P;
PR 27-JUL-1999; 99US-0145919P;
PR 28-JUL-1999; 99US-0145951P;
PR 02-AUG-1999; 99US-0146386P;
PR 02-AUG-1999; 99US-0146388P;
PR 02-AUG-1999; 99US-0146389P;
PR 04-AUG-1999; 99US-0147038P;
PR 04-AUG-1999; 99US-0147204P;
PR 04-AUG-1999; 99US-0147302P;
PR 05-AUG-1999; 99US-0147192P;
PR 05-AUG-1999; 99US-0147260P;
PR 06-AUG-1999; 99US-0147303P;
PR 06-AUG-1999; 99US-0147416P;
PR 09-AUG-1999; 99US-0147493P;
PR 09-AUG-1999; 99US-0147935P;
PR 10-AUG-1999; 99US-0148171P;
PR 11-AUG-1999; 99US-0148319P;
PR 12-AUG-1999; 99US-0148341P;
PR 13-AUG-1999; 99US-0148565P;
PR 13-AUG-1999; 99US-0148684P;
PR 16-AUG-1999; 99US-0149368P;
PR 17-AUG-1999; 99US-0149175P;
PR 18-AUG-1999; 99US-0149426P;
PR 20-AUG-1999; 99US-0149722P;
PR 20-AUG-1999; 99US-0149723P;
PR 20-AUG-1999; 99US-0149929P;
PR 23-AUG-1999; 99US-0149902P;
PR 23-AUG-1999; 99US-0149930P;
PR 25-AUG-1999; 99US-0150566P;
PR 26-AUG-1999; 99US-0150884P;
PR 27-AUG-1999; 99US-0151065P;
PR 27-AUG-1999; 99US-0151066P;
PR 27-AUG-1999; 99US-0151080P;

PR 30-AUG-1999; 99US-0151303P;
PR 31-AUG-1999; 99US-0151438P;
PR 01-SEP-1999; 99US-0151930P;
PR 07-SEP-1999; 99US-0152363P;
PR 10-SEP-1999; 99US-0153070P;
PR 13-SEP-1999; 99US-0153758P;
PR 15-SEP-1999; 99US-0154018P;
PR 16-SEP-1999; 99US-0154039P;
PR 20-SEP-1999; 99US-0154779P;
PR 22-SEP-1999; 99US-0155139P;
PR 23-SEP-1999; 99US-0155486P;
PR 24-SEP-1999; 99US-0155659P;
PR 28-SEP-1999; 99US-0156458P;
PR 29-SEP-1999; 99US-0156556P;
PR 04-OCT-1999; 99US-0157117P;
PR 05-OCT-1999; 99US-0157753P;
PR 06-OCT-1999; 99US-0157865P;
PR 07-OCT-1999; 99US-0158029P;
PR 08-OCT-1999; 99US-0158232P;
PR 12-OCT-1999; 99US-0158359P;
PR 13-OCT-1999; 99US-0159283P;
PR 13-OCT-1999; 99US-0159284P;
PR 13-OCT-1999; 99US-0159285P;
PR 14-OCT-1999; 99US-0159329P;
PR 14-OCT-1999; 99US-0159330P;
PR 14-OCT-1999; 99US-0159331P;
PR 14-OCT-1999; 99US-0159637P;
PR 14-OCT-1999; 99US-0159638P;
PR 18-OCT-1999; 99US-0159584P;
PR 21-OCT-1999; 99US-0160741P;
PR 21-OCT-1999; 99US-0160767P;
PR 21-OCT-1999; 99US-0160768P;
PR 21-OCT-1999; 99US-0160770P;
PR 21-OCT-1999; 99US-0160814P;
PR 21-OCT-1999; 99US-0160815P;
PR 22-OCT-1999; 99US-0160980P;
PR 22-OCT-1999; 99US-0160981P;
PR 22-OCT-1999; 99US-0160989P;
PR 25-OCT-1999; 99US-0161404P;
PR 25-OCT-1999; 99US-0161405P;
PR 25-OCT-1999; 99US-0161405P;
PR 25-OCT-1999; 99US-0161359P;
PR 26-OCT-1999; 99US-0161359P;
PR 26-OCT-1999; 99US-0161360P;
PR 26-OCT-1999; 99US-0161361P;
PR 28-OCT-1999; 99US-0161920P;
PR 28-OCT-1999; 99US-0161932P;
PR 28-OCT-1999; 99US-0161993P;
PR 29-OCT-1999; 99US-0162142P;

Query Match 34.3%; Score 681.6; DB 3; Length 1991;
Best Local Similarity 69.2%; Pred. No. 8.3e-191;
Matches 978; Conservative 0; Mismatches 420; Indels 15; Gaps 3;

QY 250 GGATTGGGGGATGACCTCAGCGTGTGATGAGTGGGGGGCTCGGTGTGTGGCC 319
DB 416 GGAATGCAAGAGTTGACATGTGTATGAAATTTGGCGTTCACTAGTGAATCAGA 475
QY 330 GCGAGATGCTGAGTGGCCGGCCCTATCTTGAACGTTCCCGAGAGAGCCCGGTGTC 379
DB 476 GAGAGATGAAAGAGTTGACATCTTATCTCAGCTTTCAGATGAGAGCCCGTCATT 535
QY 380 GTTCTCTGCGATGGGGAAACACCAACACCTCTCTCTGAGAGAGGAGATA 439
DB 536 GTCTATAGCAATGGGAAAGACACTAATAGCTTTTGAAGCTGAGAGAGAGCTGTT 595
QY 440 GGGTGTGAGTTATCATGTTTCTGAATGGAAGAGTGAATGTCAAAAGCCTTACAT 499
DB 556 ACTTGGGTGTCACCTAAGTGAAGATGAGATTTGAGCTTATTAAGGAATCCAT 655
QY 500 ATCAAGAGGTGATGAATTTGACT---TCCAAGNATCTGTATCAAGCCTTATGAA 556
DB 656 TTAAAGACTGCTCATGAGCTTGGAGTGAACAACGGGTATTGAAAAACACTGAAAGA 715
```

QY 557 CTGAGCACTATGTAAGAGTATCGCTATGATGAAAAGAGCTGAGCCCTAGACCAAGTAC 616
 Db 716 CTACACGAGCTTCTCAAGGACATATCAATGATGAAAGAGCTAACTTAGACCAAGGAGAC 775
 QY 617 TACCTGTTCACTTGTGAGAAATGATGATCCACAGAGATTTTTCGCTATTTGAGACAA 676
 Db 776 TACTGTTTCATTTGAGAGTGCATGTCACAGAGCTTTTCTGCGTATCTCAATAA 835
 QY 677 ATTGCTCAAGGACGAGCATGATGACGATTTGATATGTTGATTAACAACATGATGAA 736
 Db 836 ATTGGCCAAAGAGCGGTCAATATGATGATTTGATGATCGATTTATCAACCAAGAGAC 895
 QY 737 TTGCTATATGCGGATATCTTAGAGCAACCTATCTGCTGTTGCCAAGAGACTTATGG 796
 Db 896 TTACTAATGCTGATATCTTAGAGCAACATCCCTGCGTCAAAAACACTGTTGGC 955
 QY 797 GACTGATACAGGATCCAGGATACCTGTTGATGAGGTTCTTGGAGAGGCTGAGAA 856
 Db 956 GACTGAGCAAGGAAATGCAATGCTGTTGTTACTGCTACCTGGAAGGATGAGAG 1015
 QY 857 TCTGTGCTGTAATCTACTTTAGCCGAGTGTGATGATCTGACTGCTAACCATTTGCT 916
 Db 1016 TCTTGTCTATCACTACGCTGAGGTTAGGGGTGAGATTTGACAGCTAACAGATTGG 1075
 QY 917 AAAGCCTGGGACTGAGAGAAATTCAGGATAGGAAAGATGTTGATGTTGATCTTCTGT 976
 Db 1076 AAAGCGTAGAGACTGCGGAGAAATTCAGGTTTGGAAAGATGATGATGAGTTTGACTTGT 1135
 QY 977 GATCCAAATATCTACCCACATGCAAAAGACTGTTCCATCTTAATCAATTGAAAGGCCACA 1036
 Db 1136 GATCCTAAACATATATCTGAGAGCTCATCTGTTCCCTCACTTAACGTTGATAGGACACT 1195
 QY 1037 GAACCTGCTATTTTGTGCTCAGGTTTGGATCCAGCATGATGAGAGCTGCTAGAGAA 1096
 Db 1196 GAGCTGCTTACTTGTGCTCCAGGTTTGGATCCAGCTGCTATGAGGCCAGAGAAAT 1255
 QY 1097 GGTGATATTCAGTTAGGTTTAAAGATTATACAAACCTTAAAGCTCCAGAGCCCTTAT 1156
 Db 1256 GGGACATCTCTGAGAGGTTTAAAGACTGTAACAACCCAGCTCCAGAGAACTGCTATC 1315
 QY 1157 ACCAGCAAGAGACATGATGATANGCTGTGTTGATCTAATCAGATATGCTCAAGCA 1216
 Db 1316 ACCAGATCAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1369
 QY 1217 AATGCTATGTTGACATTTGAGCACTGAGCTGAGCTTGTGATGATGATGATGATGATGAT 1276
 Db 1370 AATGCTATGTTGACATTTGAGCACTGAGCTGAGCTTGTGATGATGATGATGATGATGAT 1429
 QY 1277 AGGGTATCAGGATATTTGCTATATTTGAAGATCTATGATATCTGTGATGATGATGATGAT 1336
 Db 1430 AAGGTG-----TTACCACTTTGAGATTTAGATATCGTTGATGATGATGATGATGAT 1483
 QY 1337 AGTGAAGTATGTTGCTGTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1396
 Db 1484 AGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1543
 QY 1397 CAGCAGGCAAGTGAATCTTGAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1456
 Db 1544 CAGCAGGCAAGTGAATCTTGAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1603
 QY 1457 CTACTTCAGCAGAGGAGGATATTTTCACTTATCGAAATGTTGAGAGCAATCGCTCTGATA 1516
 Db 1604 CTACTTCAGCAGAGATATCTCTCTCTCATGAGAAATGTTCAAGAAATCTCACTCAT 1663
 QY 1517 CTAGAAAAGACGAGAGCTGTGCTGAGAGAAAGTGGGTTAATGTTCAATGATCTTCGCA 1576
 Db 1664 TTAGAGAGGTTGTTCCAAAGTATTTAGAGCAATGAGAGTGAATGATGATGATGATGATGAT 1723
 QY 1577 GAGAGCTCAAGGTTAATCATGCTGCTGATAGTCCAGATAGAGATGAGATGAGATGAGATGAG 1636
 Db 1724 GGTGATCTTAAGTAAATCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1783
 QY 1637 GAAAGCCTTCATCAGGCTTCTTTGAAAGCAT 1669

Db 1784 AGGGCTCTCACTCCGCTTCTTTGAGACCAT 1816
 RESULT 6
 AAD08626
 ID AAD08626 standard; cDNA; 1658 BP.
 AC AAD08626;
 XX
 DT 04-SEP-2001 (first entry)
 DE Wheat aspartate kinase encoding wrl.pk0046.d11 (FIS) cDNA clone.
 XX
 KW Wheat; aspartate kinase; cell transformation; transgenic plant;
 KM wrl.pk0046.d11 clone; ss.
 XX
 OS Triticum aestivum.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1320
 FT /tag= a
 FT /product= "Triticum aestivum wrl.pk0046.d11 (FIS)
 FT /note= "CDS does not include start codon"
 FT /partial
 PN NO200146393-A2.
 XX
 XX 28-JUN-2001.
 XX
 PF 19-DEC-2000; 2000WO-US034396.
 XX
 PR 21-DEC-1999; 99US-0172944P.
 XX
 PA (DUPLO) DU POINT DE MEMOIRS & CO E. I.
 XX
 PI Falco SC, Famedu OO, Thorpe CJ;
 XX
 DR WPI; 2001-418057/44.
 DR P-PSDB; AAE04358.
 XX
 FT Novel nucleic acid molecules encoding aspartate kinase useful in cell
 FT transformation and transgenic plant production procedures.
 PS Claim 8; Page 59; 64pp; English.
 CC The present invention relates to nucleic acid molecules encoding
 CC aspartate kinase useful in cell transformation and transgenic plant
 CC production procedures. Nucleic acid encoding aspartate kinase is useful
 CC for producing a transgenic plant, by transforming a plant cell with a
 CC chimeric gene comprising nucleic acid encoding aspartate kinase operably
 CC linked to a regulatory sequence and regenerating a plant from the
 CC transformed plant cell. The present sequence is cholic.pk002.k6 (FIS) cDNA
 CC clone encoding Triticum aestivum aspartate kinase
 SC Sequence 1658 BP; 457 A; 331 C; 417 G; 453 T; 0 U; 0 Other;
 Query Match 34.0%; Score 674.6; DB 5; Length 1658;
 Best Local Similarity 72.8%; Pred. No. 8.8e-189;
 Matches 930; Conservative 0; Mismatches 326; Indels 21; Gaps 4;
 QY 425 GAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 484
 Db 4 GAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 63
 QY 485 GTCAAAAGCCTACATATCAAGACGAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 544
 Db 64 ATCAAGAGACTCATCTCAGACCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 123
 QY 545 AGCCTTT---ATGAAGTGAAGCACTATTTGAAGTATCGTATGATGATGATGATGATGATGATGAT 601
 Db 124 GGTTTTGTGAGAGAGTGAAGCACTGCTCAAGGAGTGTGCTATGATGATGATGATGATGATGATGAT 183

QY 602 CCTAGGACGAGTACTACTTGTTCATTTGGAGATGATGTCACACGAGATTTTTC 661
DB 184 CTTAGGACAGGATTAACCTTGTTCCTTTGGTGAATGATGTCACAGAAATATTTTC 243
QY 662 GCTTATTTGAACAAATTCGTGTCAAGGACGGGAGTGAACGATTTGATTTGTTCC 721
DB 244 GCATTTTGAATTAATCTAGGAGAGAGGACGACGATTAATGCTTTGATCTTGCTTT 303
QY 722 ATTACACTGATGAATTTGTATGCGATATCTTAAAGCAACCTATCTGCTGTGCG 781
DB 304 ATAAACCACTGACGATTTCAAAATGCCGATATCTCGAAGCAACTTATCCTGCTGTGCA 363
QY 782 AAGAGCTTCATGGGACCTGATACAGATTCAGGAGTAACTGTTTACTGGGTCTT 841
DB 364 AAGAGGCTACATGAGATTTGATGATGACCTGCTATTTCTTAATGATGCTGTTCTT 423
QY 842 GGGAGGGCTGGAATCTGTGCTGTAATCTTAATCTTTAGCCGAGGTGTAGTACTGACT 901
DB 424 GGGAGGGATGGAATCTGTGGGTCAACACGTTAGAGAGGGGCGGAGTGACTTGACC 483
QY 902 GCTACACCACTTGTGTAAACCTTTGGACCTGAGAGAAATTCAGTATGGAAGATTTGAT 961
DB 484 GCTACACCACTTGTGTAAACCTTTGGAGTAAATTCAGGTTTGAAGAGATGTAGAC 543
QY 962 GGTGTACTTACTTGTGATCCAAATATCTACCCACATGCAAGACTGTTCCATCTTACA 1021
DB 544 GGTGTGTACGTTGATCCAAATATTTATGCAACGCGGTACACGATCCTACTTGACT 603
QY 1022 TTTGAAGAGCCACAGAACTTGTATTTTGTGTCTGAGTTTGGATCCCAATCGATG 1081
DB 604 TTTGATGAGCAGCTGAACTTGTATTTTGTGACAGGTTTGGATCCCAATCCATG 663
QY 1082 AGACCTGTAGAGAAAGTATATTCAGTTAGGGTTAAGAAATTCACAACTTAAGCT 1141
DB 664 CGACCAAGGAGGAGGTGTATCCAGTTCAGAGTGAAGAACTCATAAACGTCATGCA 723
QY 1142 CCAGGACACCTTATTAACAAGAAAGACATGATTAANGTCTGCTGTCTACTAGC 1201
DB 724 CCGGACACTGTGATCTATAAACAGAGATATGCCCAAG-----CATATTAACAGC 777
QY 1202 ATAGCTCAAGCAATGCAATGCAATGTTGACATTTGAGACTGTGAGACTCGATCTGTGAC 1261
DB 778 ATTTGCTGAATCAAAATTTACATGCTGATTAATGAGACAGAGATGCTGACAG 837
QY 1262 TATGTTTTCTGCAAGGGTATCAGGATTTGCTATATTGAAGATCTATGATATCTGTG 1321
DB 838 TATGCTTTCTAGCAAGGCTTCTCAATAT-----TTGAAGATTTGGGTATCTGTGTT 891
QY 1322 GATTTGTTGTACCAAGTGAAGTGTGTTCTGTGATCACTTGATCCATCAAGTCTGG 1381
DB 892 GATTTGTTGTCTATGTAAGTCAATATCAATGATCACTGATCCATCAAGTCTGG 951
QY 1382 AGTAGGAACTGATACAGAGGCAAGTGAACCTGACCATGATTTGAAGACTTGAGAA 1441
DB 952 AGTGTGAATGATTCAGAG-----GAGCTTGATCAATGATTTGAAGACTTGAGAA 1005
QY 1442 ATAGCAATTTGTTGTTACTTACAGAGAGGCGATTAATTTCTTATGGAAATGTGAG 1501
DB 1006 ATTTGCGTTGTTATCTCTACAGCAATCAATCAATTTCCCTATGAGGAAATGTGAG 1065
QY 1502 CAATGCTCTGATCTAGAAAAGACGGGAGTGTCTGAGGAAAGTGGGTTATATTT 1561
DB 1066 AGATCGCTCTGATTTCTTGAGAAAGCGGTTCAATGTTCTACGAGAAATGAGTTATTT 1125
QY 1562 CAATGATCTCGAAGAGAGGCTCAAGGTTAATGATGCTGATAGTCCATGATAGGAT 1621
DB 1126 CAATGATTTTGAAGAGGCGCTCAAGGATCACTCTGTTGTTGAATGACGACGAG 1185
QY 1622 GCAAGGCACTGCTAGAGCCCTTCAATCAAGGCTTTTGAAGACGATGCTCATCA 1681
DB 1186 GCGAAGGAGTGGCGTGAAGCCCTTCACTGCGGCTTTTGAAGACGTTTCTTGTCA 1245

QY 1682 GTCCAGCGCGAGACT 1698
DB 1246 GTAGGAGAGCGGACCT 1262
RESULT 7
AAC41888
ID AAC41888 standard; DNA; 1395 BP.
XX AAC41888;
XX
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 33506.
XX
XX
KW Hybridization assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX
XX EP1033405-72.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132464P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132566P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137502P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.

```
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139762P.
PR 21-JUN-1999; 99US-0138117P.
PR 22-JUN-1999; 99US-0138999P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140931P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142777P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144353P.
PR 20-JUL-1999; 99US-0144354P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145318P.
PR 27-JUL-1999; 99US-0145319P.
PR 28-JUL-1999; 99US-0145351P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147202P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.

PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149729P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149903P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151067P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158293P.
PR 13-OCT-1999; 99US-0158294P.
PR 13-OCT-1999; 99US-0158325P.
PR 14-OCT-1999; 99US-0158328P.
PR 14-OCT-1999; 99US-0158330P.
PR 14-OCT-1999; 99US-0158331P.
PR 14-OCT-1999; 99US-0158637P.
PR 14-OCT-1999; 99US-0158638P.
PR 18-OCT-1999; 99US-0158584P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160880P.
PR 22-OCT-1999; 99US-0160981P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161362P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161932P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 33 1%; Score 656.2; DB 3; Length 1995;
Best Local Similarity 69.0%; Pred. No. 2.8e-183;
Matches 975; Conservative 0; Mismatches 420; Indels 18; Gaps 5;
```

QY 380 GTTCTCTGCGATGGGGAAAAACCAACCACTTCTCTGCTGGAGAGAGCAGTA 439
 Db 541 GTGCTATACGAAATGGGAAAAGACATATATAGCTTTTGAAGCTGGAGAGAGCTGT 600
 QY 440 GGGGTGGAGTTATCCATGTTTCTGAATCGAAGATGGAATATGCTCAAGCCCTACAT 499
 Db 601 ACTTCGGGTGCATCAACGTTAGAAAGTATGAAATTTAGCTTTTAAAGAACTCCAT 660
 QY 500 ATCAGAGCGATGAACTTGGACT--TCCAGNATCTGTATATCAAGCTTTATGAA 556
 Db 661 TTAAAGACTGCTCATAGCTTTGAGTGAACCAACGCTTATGAAAACCTTAGAAGA 720
 QY 557 CTGAGCAACTATTTGAAAAGTATGCTATGATGAAAGAGCTGACGCTTAGACAGTAC 616
 Db 721 CTACACCAAGCTTCTCAAGGATATCATATGATGAAAGAGCTTAACCTTACGACAGAGGAC 780
 QY 617 TACCTTTTTCATTTTGGAGAAATGCATGTCACCGAGATTTTCTGCTTATTTGAACAA 676
 Db 781 TACTTGTTCATTTGGAGAGTGCATGTCACCAAGGCTTTTCTTGGTATCTCATTA 840
 QY 677 ATTGCTGCAAGGCAAGGAGATGACGCAATTTGATATTTGTTTCACTTACACTGATGA 736
 Db 841 ATTGGCCAAAGGCGCTCAATATGATGATTTGATGATCGGATTTATCACACAGACGAC 900
 QY 737 TTTGGTAATGGGATATCTTGAAGCAACCTATCTGCTGTTGGAGAGACTTCATGGG 796
 Db 901 TTTACTAATGCTGATATCTTGAAGCAACATACCTGCGCTCAAAAACACTTGTGGC 960
 QY 797 GACTGATACAGATCCAGCATCCTGTTGTAAGTCTGTTGTAAGGAGGCTGAAA 856
 Db 961 GACTGAGCAAGGAAATGCACTCCCTGTTGTAAGTCTGTAAGGAGGAGG 1020
 QY 857 TCTGCTGCTAATCTTACCTTTAGGCGGAGGTGTAGTGACTGATCTTAACCATTTGT 916
 Db 1021 TCTTGTGCTATCACTACCTGCTGGTGAAGGTGAGTGTGACAGCTACACCATTTGG 1080
 QY 917 AAAGCCTGGAGCTAGAGAAATTCAGGTATGAAAAGATTTGATGCTTACTTGT 976
 Db 1081 AAAGGCTAGAGCTGCGGAAATTCAGGTTTGGAAAGATGATGAGTTTGAATGT 1140
 QY 977 GATCCAAATATCTACCCATGCAAGCAAGCTGTTCCATCTTAACATTTGAAGAGCCACA 1036
 Db 1141 GATCTTAATATATCTGAGACTCAATCTGTTCCGTACTTAAGTTCATGAGGAGCT 1200
 QY 1037 GAACCTGCTATTTTGGGCTCAGGTTTGCATCCACATGATGAGAGCTGTAGAGAA 1096
 Db 1201 GAGCTTGCTTACTTTGGTCCAGAGTGTGCATCCATCTGTCTTAGAGCCAGC-ATAGAT 1259
 QY 1097 GGTGATATTCAGATTAGGTTAAGATTCATCAACCTTAAAGCTCCAGGACCTTATT 1156
 Db 1260 GGGGACATTCCTGTAGGGTTAAGAACTCGTACAACTCCAGCTCCAGGAACTGTATC 1319
 QY 1157 ACCAGCAAGAGACATGATATANGTCTGTTGATACATATAGATATGTCCTCAAGTCA 1216
 Db 1320 ACCAGATCAAGAGACATGATATAGGCTTTG-----CTAACGACATTTCTTAAGAGT 1373
 QY 1217 AATGTCATATGTTGACATTTGAGCACTCGATGCTTGTGCTAGTATGTTTCTGGCA 1276
 Db 1374 AATGTCATATGTTGACATTTGAGCACTCGATGCTTGTGCTAGTATGTTTCTGGCA 1433
 QY 1277 AAGGTATCAAGTATTTGCTATATTTGAAGTCTATATATCTGAGATGTTGTTGCTACC 1336
 Db 1434 AAGGTATTTTACCACTTTGAGATTTAGGATATGCTTTGATGTTGTTGCTAACCA 1487
 QY 1337 AGTGAAGTATGTTTCTGTCATCTGATTCATCAAAATCTGAGATGAGAACTGATA 1396
 Db 1488 AGTGAAGTATGTTTCTGTCATCTGATTCATCAAAATCTGAGATGAGAACTGATA 1545
 QY 1397 CAGCAGCAAGTGAATTTGATCACTGATGATGAGAGCTTGAAGAAATAGCAATTTGTTGT 1456
 Db 1546 CAGCAGCAAGTGAATTTGATCACTGATGATGAGAGCTTGAAGAAATAGCAATTTGTTGT 1605

QY 1457 CTACTTCAGAGAGGGCGATATTTCTACTTATTCGGAAAATGTGAGCAATTCCTCTGTATA 1516
 Db 1606 CTACTTCAGAGAGATCAATCATCTCTCATAGGAAAATGTCAAAAATCTCACTCATTA 1665
 QY 1517 CTGAAGAAAGCGGAGCGTGTGCTGAGAGAAAAGTGGGTTAATGTTCAAGTATCTCGCA 1576
 Db 1666 TTAGAGAGGAGTGTTCAGAGATTTAGAACCAATGAGATGATGCAATGATCTCGCAG 1725
 QY 1577 GGAGCGTCAAGGTTAATCATGCTGATGATGCTGATGATGATGATGATGATGATGAT 1636
 Db 1726 GGTGATCTTAAGGTTAAACATCTCATTTGATGATGATGATGATGATGATGATGATG 1785
 QY 1637 GAGGCTTCATCAGGCTTCTTTGAAGCAT 1669
 Db 1786 AGGCTCTCCATCTCCGCTTCTTTGAGATCAT 1818
 RESULT 8
 AAC32977
 ID AAC32977 standard; DNA; 2164 BP.
 XX AAC32977;
 AC 17-OCT-2000 (first entry)
 XX
 DT
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 1340.
 XX
 KM Hybridisation assay; genetic mapping; gene expression control;
 KM protein identification; signal transduction pathway; metabolic pathway;
 KM promoter; termination sequence; ss.
 OS Arabidopsis thaliana.
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-00301439.
 XX
 PR 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 23-APR-1999; 99US-0130891P.
 PR 28-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 04-MAY-1999; 99US-0132404P.
 PR 05-MAY-1999; 99US-0132486P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 06-MAY-1999; 99US-0132487P.
 PR 07-MAY-1999; 99US-0132863P.
 PR 11-MAY-1999; 99US-0134256P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 14-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.
 PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0136293P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 16-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140655P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141827P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142054P.
PR 06-JUL-1999; 99US-0142290P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144844P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.

PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 26-AUG-1999; 99US-0150565P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-015659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156566P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 29-OCT-1999; 99US-0162424P.

Query Match 33.0%; Score 654.6; DB 3; Length 2164;
 Best Local Similarity 68.9%; Pred. No. 8,9e-183;
 Matches 964; Conservative 0; Mismatches 415; Indels 21; Gaps 4;

266 GGGGATCAGCTCAGGCTGTGTATGAGTTCGAGGGGCTCTCGGTGTGCGGCGGAG 325
 Db GAGAAAAATTGACTGTGTATGAAATTCGCTGCTATCAGTGCATCAGCAAGAG 618
 QY 326 ATGCTGAGGTGGCGGCTCATCTGACGTTCCCGAGAGAGCCCGCTGTCTTC 385
 Db ATGATACAAAGTTGCCAAACTCACTACGTTCCAGATGAGAAAGCTGTGTGTGCTA 678
 QY 386 TTGCGCATGGGAAAAACCAACAACCTTCTCTGCTGTGAGAGAGAGGAGTGT 445
 Db TTAGCAATGGCAAGACCAACATAGCTTTTGTATGCTGTGGGAGAGGCTGTGCTGC 728
 QY 446 GGAGTTATCCATGTTTCTGAAATCGAAGATGGAATATGATCAAAAGCTTACATATCAAG 505
 Db GGATGACCAACGTGACACTATCGAAGGTTGAGCTATATPAAAGAACTCCATATPAGG 798
 QY 506 ACCGTGATGACTTGACTTCCAGATTCGTGATAC--AAGCTTATGATGAGAG 562
 Db ACTGCTCATGAGCTTGGATGGAAACAGCAGTATTGCAAACTTGAAGAGCTGGA 858
 QY 563 CAACTATTGAAAGATATGCTATGATGAAAGAGCTGACGCTTAGAGACAGTACTACTT 622
 Db CAGCTTCTTAAAGGGGTGCGCATGATGAAAGAAATTAATTACCTGAAAGACTACTTG 918
 QY 623 GTTTCATTGGAAATGCATGTCACCGAGTTTCTGCTTATTGTAACAAATTCGT 682
 Db GTTTCATTGGAAATGATGTCACCTGAGCTATTGCTGTGTCCTGAAATAAATTCGT 978
 QY 683 GTCAAGGCAAGGAGATGACGATTTGATATGTTGTTTCACTTAAACATGATGATTTGT 742
 Db CACAAAGCAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1038
 QY 743 AATGCGATATCTTGAAGCAACCTATCTGCTGTGTGGAAGAGACTTCACTGGGCACTGG 802
 Db AATGCTGATATCTTGAAGCAACATACCGGCGGTTTGAAAGAAATTAATTGCTGATCGG 1098
 QY 803 AATAGGATCCAGCACTGTTGTTTCTGCTGCTGCTGGAAGAGGCTGGAATTCGT 862
 Db ACGAAGAAAAATGCACTCCCGCTGTGTCAGAGCTTCTTGAAAGAGATGAGATCTTGT 1158
 QY 863 GCTGTAATCTTAAAGCCGAGGTGTATGATGATGATGATGATGATGATGATGATGATGAT 922
 Db GCTGTAATCTTAAAGCCGAGGTGTATGATGATGATGATGATGATGATGATGATGATGAT 1218
 QY 923 TTGGACATGAGAAATTAAGGATGAGAAAGATGATGATGATGATGATGATGATGATGATGAT 982
 Db TTAGGTTTGGGAGATCCAGGTTTGAAGAGATGATGATGATGATGATGATGATGATGATGAT 1278
 QY 983 AATATCTACCAATGCAAGAAGCTGTTCTATCTTAACTTGAAGAGCCACAGAACTT 1042
 Db AACATTTACGCGAGACTCAACCTGTTCCACCTTGAATGATGATGATGATGATGATGATGAT 1338
 QY 1043 GCTTATTTGGTGTCAAGTTTGTGATCCCAATGAGATGAGAACCTGCTGAGAGAGTGT 1102
 Db GCTTATTTGGTGTCAAGTTTGTGATCCCAATGAGATGAGAACCTGCTGAGAGAGTGT 1398
 QY 1103 ATTCAAGTTAGGTTTAAATTCATACAAACCTTAAAGCTTCAAGGACCTTATTACAGA 1162
 Db ATTCTGTAAAGGTTTAAAGCTTAAATTCATACAAACCTTCAAGGAACTGATCACCAGA 1458
 QY 1163 CAAAGACATGATATANGGTCTGTGTGATTAATTAACATAGTCTCAATCAATGTC 1222
 Db TCAAGACATGATGCA--GCTGTATGACACAGATGCTTCTAAAGCTAATGTC 1512
 QY 1223 ACTATGTTGACATTTGTGACACTCGAGTGTGTGCTGATGTTTCTGTGCAAGGTA 1282
 Db ACCATGTTGACATCTAGCAACCGGTATGCTCGGCAATATGTTTCTGTGCAAGGTG 1572

QY 1283 TCAGTATTTGCTATATTGAAGATCTATATATCTGTGATGTTGTTCTACAGTGA 1342
 Db 1573 T-----TTTCCATTTGAAAAAGTTGGCATCTGTGTGATCTTTGTCAACAGCGAA 1626
 QY 1343 GTTATGTTTCTGTGCTACTGATTCATCAAGATCTGAGTGGGAATGATACAGCAG 1402
 Db 1627 GTTATGATTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1685
 QY 1403 GCAAGTAACTTGCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1462
 Db 1686 -----CGAGCTTGTATCGAGTGTGAGAGAACTTGAAGAAATGCTTGTGAATCTGCTT 1740
 QY 1463 CAGAGAGGCGGATTAATTTCACTTATGGAATGTGAGCAATGCTCTGATCTAGAA 1522
 Db 1741 CGACAGATCAATTAATCTCTCATGCGAAATGTTCAAGATCATCATTAATATTAGAG 1800
 QY 1523 AAGAGGAGCTGTGCTGAGAGAAAGTGGGTTAATGTTCAATGATGATGATGATGATGATGATGAT 1582
 Db 1801 AAGGCTTCCAGATCTTCAACCAATGAGATTAATGTCAGATGATCTCCAGAGTGA 1860
 QY 1583 TCAAGGTTAATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1642
 Db 1861 TCTAAGTAAACATCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
 QY 1643 CTTCATCAGGCTTCTTGA 1662
 Db 1921 CTCACCTGCGCTTCTTGA 1940

RESULT 9
 AAC48615
 ID AAC48615 standard; DNA; 1692 BP.
 AC AAC48615;
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 58119.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway; metabolic pathway;
 KW promoter; termination sequence; ss.
 OS Arabidopsis thaliana.
 XX
 PN EP103405-A2.
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000BP-00301439.
 XX
 PR 25-MAR-1999; 99US-0121825P.
 PR 05-MAR-1999; 98US-0133180P.
 PR 09-MAR-1999; 98US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 98US-0128724P.
 PR 16-APR-1999; 98US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130044P.
 PR 23-APR-1999; 99US-0130510P.
 PR 23-APR-1999; 99US-0130891P.
 PR 28-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 30-APR-1999; 98US-0132407P.
 PR 04-MAY-1999; 98US-0132484P.
 PR 05-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 06-MAY-1999; 99US-0132487P.
 PR 07-MAY-1999; 99US-0132863P.

[illegible]

PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161933P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 31.0%; Score 616.2; DB 3; Length 1692;
 Best Local Similarity 66.6%; Pred. No. 1.9e-171;
 Matches 979; Conservative 0; Mismatches 419; Indels 72; Gaps 4;

260 GGATTGGGGGATCAGCTGAGTGTGATGAGTGTGGGGGCTGCTGGTGTGCTGGGC 319
 229 GGAATGGCAAGAGTTGACATGTGTAAATTTGGCGTTTATCAGTGAATCAGCA 288
 320 GCGAGATGCTGAGTGGCCGCTCATCTGACGTTCCCGAGAGCGCCCGTGTGTC 379
 289 GAGAGATGAAAGAGTTGCCATCTTATCTCAGCTTCCAGATGAGAGCCGTCATT 348
 380 GTTCTCTGCGCATGGGGAAAACGACCAACACTTCTCTGCTGGAGAGGAGCA 439
 349 GTGCTATCAGCAATGGGAAAGCAGCTAATGACTTTGAAAGCTGGAGAAAGCTGTT 408
 440 GGGTGTGAGTATTCATGTTTCTGAAATCGAAGTGAATATGTCAAAAGCTCAT 499
 409 ACTTGGCGGTGCTACATGATGAAAGTATGAAATGAGCTTTAAAGAACTCCAT 468
 500 A-----TC 502
 469 TTAAGTGTCAAGTCCAAATTAGCTTATCTCATGGAATAAATTTCTATGTTTC 528
 503 AAGCGGTGATGAACTTGAATTCGAAGNATCTGTAAATCAAG---CCTTTATGAACTG 559
 529 AGAATGCTCATGAGCTTGGAGTGAACAAAGCTTATGAAAACACCTGAAAGAGCTA 588
 560 GAGCACTATTGAAAGATATGCTATATGAAAAGCTGACGCTTAGAGCACTGCTAC 619
 589 CACGAGCTTCCAAAGGATATCAATGATGAAAGCTAACCTTAGCACAAGGAGCTAC 648
 620 CTTGTTTCATTTGAGATGATGATGCCACGAGATTTTCTGCTTATTTGAACAATTT 679
 649 TTGTTTCATTTGAGAGTGCATGTCACAAAGCTTTCTCTGGTATCTCAATAAATTT 708
 680 CGTGTCAAGCAGCGCAGTATGACGATTTGATTTGTTTCAATACACTGATGATTT 739
 709 GAGCACAAGGCGCTCATATGATGCAATGAGATCGAATTTATCCACACAGCGACTTT 768
 740 GATATGCGGATATCTTGAAGCAACCTATCTGCTGCGAAGAGACTTCAATGGGAGC 799
 769 ACTATGCTGATATCTTGAAGCAACATACCTGCGCTCCAAAACACTGTGTGGCGAC 828
 800 TGGATACAGATCCAGCGATACCTGTTGTTACTGAGTTCTTGGGAAAGGCTGGAATCT 859
 829 TGGAGCAAGGAAATGCGATCCCTGTTGTTACTGCGTACCTCGAAAGGAGTGAAGTCT 888
 860 GATGCTTAACACTTATGAGCGAGTGTGATGATGCTGCTACCAACCACTTGGTAA 919
 889 TGTGTATCATCACTGCTGGTGAAGGAGTGTGAGTATTTGACAGTACACGATTTGGGAAA 948
 920 GCTTGGGACTGAGAGAAATTCAGGTATGAAAGATGTTATGTTGTTACTTCTGTAT 979
 949 GCGTAGGAGCTGCGGGAATTCAGGTTTGAAGAAATGATGAGTGGATTTTGAACCTTGAT 1008
 980 CCAATATCTCCCAATGCAAGAGCTGTTCCATACTTAACTTGAAGAGCCACAGAA 1039
 1009 CCTAACATATATCTCGAGCTCAATCTGTTCCGACTTAAAGCTTCATGATGAGGAGCTGAG 1068

QY 1040 CTTGCTATTTTGTGCTCAGGTTTGCATCCAAATGATGAGCTGTGAGAGAGT 1099
 DB CTTGCTATCTTGTGCTGCCAGGTTGTGCATCCAGTGTATGAGCCGAGCAAGATGGC 1128
 QY 1100 GATTTTCATTTAGGTTTGAAGATTCATCAACCTTAAAGCTCCAGGCACTTTTAC 1159
 DB GACTTCTCTGAGGTTTGAAGCTGTACAAACCTGCTCCAGGAACTGTCTATCCAG 1188
 QY 1160 AGACAAAGACATGATTAAGTCTGTTGTTACTAATGATGATGATGATGATGAT 1219
 DB AGATCAAGACATGATTA-----GCTGTGTACTAACGAGATGTTCTGAAGATAT 1242
 QY 1220 GTCACTATGTTGACATTTGAGCACTCGGATGCTTGTGCTGATATGTTTCTGGCAAG 1279
 DB 1243 GTGACCATGTTGACATGACAGACACCTGATGCTTGGCCAAATGATGTTCTTCCCAAG 1302
 QY 1280 GTATCAGTATTTGCTATATTTGAAGATCTATGATATATCTGATGATGTTGCTTACAGT 1339
 DB 1303 GTGT-----TTACCAATTTGAAAGTTTAAAGTATGCGTTGATGTTGTGCAACAGT 1356
 QY 1340 GAAGTTATGTTTCTGTGCTATTCATCATCAAGATCTGAGTGAAGGAACTGATACAG 1399
 DB 1357 GAAGTTATGATATCATATGATGATCCAGAAAGCTTGGGTTAGAGATTAATGAG 1416
 QY 1400 CAGCAAGTGAACCTGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1459
 DB 1417 CGATTAATGATATGATATTTTGTGAGAGCTAGAGAAATGCTGTTGTGAACCTA 1476
 QY 1460 CTTGACGAGAGGCGATATTTTCACTTATGCGAATGTGAGCAATCTGTTCTGATCTA 1519
 DB 1477 CTTCAAGCAGATCAATCTCTCTCATAGGAAATGTTCAAAATCTCTCACTCATATTA 1536
 QY 1520 GAAAGAGCGGAGCTGTGCTGAGAGAAAGTGGGTTAATGTCAGATGATCTCCAGAGA 1579
 DB 1537 GAGAGGTTTCCAAATTTTGAAGACATGAGAGTGAATGTCAGATGATCTCGAGAGGT 1596
 QY 1580 GCGTAAAGGTTTAAATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1639
 DB 1597 GCATCTGAGTAAATCTCATATGATGATGATGATGATGATGATGATGATGATGATGAT 1656
 QY 1640 GCGCTCATGAGCGTCTTTTGAAGAGAT 1669
 DB 1657 GCTTCCATCTCGCCTCTTTGAGACGAT 1686

RESULT 10
 AAD08619
 ID AAD08619 standard; cDNA; 565 BP.
 XX
 AC AAD08619;
 DT 04-SEP-2001 (first entry)
 XX
 DE Zea mays aspartate kinase encoding bmsl.pK0008.e5 cDNA clone.
 XX
 DE Zea mays aspartate kinase; cell transformation; transgenic plant;
 KW bmsl.pK0008.e5 clone; ss.
 KW
 OS Zea mays.
 XX
 FH Key
 XX CDS
 FT 2..295 location/Qualifiers
 FT /*tag= a
 FT /product= "Zea mays bmsl.pK0008.e5 aspartate kinase"
 FT /transl_except= (pos:125..127, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid; CDS
 FT does not include start codon"
 FT /partial
 XX
 PN MO200146393-A2.
 XX
 PD 28-JUN-2001.

PF 19-DEC-2000; 2000WO-US034396.
 XX
 PR 21-DEC-1999; 99US-0172944P.
 XX
 XX (DUPO) DU POINT DE NEMOURS & CO E I.
 XX
 PI Falco SC, Famodu OO, Thorpe CJ;
 XX
 XX WPI: 2001-418057/44.
 DR P-PSDB; AAE04351.
 XX
 XX Novel nucleic acid molecules encoding aspartate kinase useful in cell
 PT transformation and transgenic plant production procedures.
 XX
 XX Claim 8; Page 47; 64pp; English.
 XX
 CC The present invention relates to nucleic acid molecules encoding
 CC aspartate kinase useful in cell transformation and transgenic plant
 CC production procedures. Nucleic acid encoding aspartate kinase is useful
 CC for producing a transgenic plant, by transforming a plant cell with a
 CC chimeric gene comprising nucleic acid encoding aspartate kinase operably
 CC linked to a regulatory sequence and regenerating a plant from the
 CC transformed plant cell. The present sequence is bmsl.pk0008.es cDNA clone
 CC encoding Zea mays aspartate kinase
 XX
 SQ Sequence 565 BP; 158 A; 115 C; 137 G; 154 T; 0 U; 1 Other;

Query Match 28.0%; Score 555; DB 5; Length 565;
 Best Local Similarity 98.9%; Pred. No. 1.4e-153;
 Matches 558; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1418 CATGATGTTGAAGACCTGAGAAATAGCAATTTGCTCTTCACTTTCAGAGAGCGCATA 1477
 DB 2 CACGAGGTTGAAGACCTGAGAAATAGCAATTTGCTCTTCACTTTCAGAGAGCGCATA 61
 QY 1478 ATTTCACCTATCGAAGAAATGTGAGCAATCGTCTCTGATCTAGAAAAGACGCGACGTGTG 1537
 DB 62 ATTTCACCTATCGAAGAAATGTGAGCAATCGTCTCTGATCTAGAAAAGACGCGACGTGTG 121
 QY 1538 CTGAGAAAAGTGGGGTTATGTTTCAATGATCTGCGAAGAGCGTCAAGAGTTACATG 1597
 DB 122 CTGAGAAAAGTGGGGTTATGTTTCAATGATCTGCGAAGAGCGTCAAGAGTTACATG 181
 QY 1598 TCGCTGATGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1657
 DB 182 TCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 241
 QY 1658 TTTGAAGACGATGCTCTATCACAAGTCGAAGCGAAGAACTGATCTGCGGCTGATCAACG 1717
 DB 242 TTTGAAGACGATGCTCTATCACAAGTCGAAGCGAAGAACTGATCTGCGGCTGATCAACG 301
 QY 1718 TAGGCTTTGCTGGGTCGAGGCGGTATCTGTTAGATTCCTGCCCTCCATGAAACG 1777
 DB 302 TAGGCTTTGCTGGGTCGAGGCGGTATCTGTTAGATTCCTGCCCTCCATGAAACG 361
 QY 1778 GCATGGGCAATTTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1837
 DB 362 GCATGGGCAATTTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 421
 QY 1838 AACCAATGATCTGCAAAACCTGCTTCTTTTGAAGCTGTTGAGACACAGAGACGCTG 1897
 DB 422 AACCAATGATCTGCAAAACCTGCTTCTTTTGAAGCTGTTGAGACACAGAGACGCTG 481
 QY 1898 CGAGACCGGATGTCACAAAGATGCAAGCTGTCGATATTAATTCACATGCTCTTAATCA 1957
 DB 482 CGAGACCGGATGTCACAAAGATGCAAGCTGTCGATATTAATTCACATGCTCTTAATCA 541
 QY 1958 AT 1981
 DB 542 AT 565

RESULT 11

AAD06620
 ID AAD06620 standard; cDNA; 513 BP.
 XX
 AC AAD06620;
 XX
 XX 04-SEP-2001 (first entry)
 XX
 DE Zea mays aspartate kinase encoding choic.pk002.k6 cDNA clone.
 XX
 KW Corn; aspartate kinase; cell transformation; transgenic plant;
 XX choic.pk002.k6 clone; ss.
 XX
 OS Zea mays.
 XX
 XX Key Location/Qualifiers
 FH 1..459
 FT CDS
 FT /*tag= a
 FT /product= "Zea mays choic.pk002.k6 aspartate kinase"
 FT /note= "CDS does not include start codon"
 FT /partial
 XX
 PN WO200146393-A2.
 XX
 PD 28-JUN-2001.
 XX
 PF 19-DEC-2000; 2000WO-US034396.
 XX
 PR 21-DEC-1999; 99US-0172944P.
 XX
 XX (DUPO) DU POINT DE NEMOURS & CO E I.
 XX
 PI Falco SC, Famodu OO, Thorpe CJ;
 XX
 XX WPI: 2001-418057/44.
 DR P-PSDB; AAE04352.
 XX
 XX Novel nucleic acid molecules encoding aspartate kinase useful in cell
 PT transformation and transgenic plant production procedures.
 XX
 XX Claim 8; Page 48; 64pp; English.
 XX
 CC The present invention relates to nucleic acid molecules encoding
 CC aspartate kinase useful in cell transformation and transgenic plant
 CC production procedures. Nucleic acid encoding aspartate kinase is useful
 CC for producing a transgenic plant, by transforming a plant cell with a
 CC chimeric gene comprising nucleic acid encoding aspartate kinase operably
 CC linked to a regulatory sequence and regenerating a plant from the
 CC transformed plant cell. The present sequence is choic.pk002.k6 cDNA clone
 CC encoding Zea mays aspartate kinase
 XX
 SQ Sequence 513 BP; 102 A; 140 C; 170 G; 100 T; 0 U; 1 Other;

Query Match 22.6%; Score 448.6; DB 5; Length 513;
 Best Local Similarity 96.6%; Pred. No. 4.6e-122;
 Matches 479; Conservative 0; Mismatches 15; Indels 2; Gaps 2;

QY 8 ACAGACGAGAGGATCTAGAAATGCAATCCAGTGGATCGGCTGCGCGCCCGCCG 67
 DB 1 ACAGACGAGAGGATCTAGAAATGCAATCCAGTGGATCGGCTGCGCGCCCGCCG 60
 QY 68 CTCGTTCCGTGATACCTCCGCGGAGCTCTGACATGTTGAGAGACTGCGGCTTCG 127
 DB 61 CTCGTTCCGTGATACCTCCGCGGAGCTCTGACATGTTGAGAGACTGCGGCTTCG 120
 QY 128 ACCCGAACCGGCGCTCGCGGTGCAAGAGGTTGTCATGATGATGATGATGATGATGATG 187
 DB 121 ACCCGAACCGGCGCTCGCGGTGCAAGAGGTTGTCATGATGATGATGATGATGATGATG 180
 QY 188 CGTGGGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 247
 DB 181 CGTGGGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
 QY 248 CTCGGGATGAGGAGGATGCGGAGATCACTAGCGTGTGATGAAGTTGCGGCGGCTCTCG 307

Accession	Sequence	Length
Dh	CTGGGATGAGAGGAAATGGAGGATCAGCTCAGCAGTGAGTGAATTTGGGGGGTCTCTCG	300
Dh	241	
Qy	GTGTGTCGGCCCGAGAGATGGCTGAGTGGCCGGCCCTCATCTGACCTTCCCGAGAG	367
Dh	308	
Dh	GTGTGCTGGCCCGAGAGATGGCTGAGTGGCCGGCCCTCATCTGACCTTCCCGAGAG	360
Qy	301	
Qy	CGCCCGTGGTCTTCTCTGTCGATGGGGGAAACACCAACAACCTTCTCTGCT-GG	426
Dh	368	
Dh	CGCCCGTGGTCTTCTCTGTCGATGGGGGAAACACCAACAACCTTCTCTGCTGGG	420
Qy	361	
Qy	AGAGAAAGCAGTGGGATGAGATTATCCATCTTTCTGAAAT-CGAAGAGTGAATGG	485
Dh	427	
Dh	AGAGAAAGCAGTGAAGTGTGAGATTATCTGTTTCTGAATCCGAGAAATGNAATGG	480
Qy	421	
Qy	TCAAAAGCCTACATAT 501	
Dh	486	
Dh	TCAAAAGCCTAAAT 496	

RESULT 12

ID	AAD08624	standard; cDNA; 847 BP.
XX	AAD08624:	
AC		
XX		
DT	04-SEP-2001	(first entry)
DE	Oryza sativa aspartate kinase encoding rdrlf.pk005.f20 (FIS) cDNA clone	
XX		
KM	Rice; aspartate kinase; cell transformation; transgenic plant;	
KM	rdrlf.pk005.f20 clone; ss.	
OS	Oryza sativa.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	3..845
FT	/*tag=	a
FT	/product=	"Oryza sativa rdrlf.pk005.f20 (FIS) aspartate
FT	kinase"	
FT	/note=	"CDS does not include stop codon"
FT	/partial	
XX		
XX	MO200146393-A2.	
PD	28-JUN-2001.	
XX		
PF	19-DEC-2000; 2000MO-US034356.	
XX		
PR	21-DEC-1999; 99US-0172944P.	
XX		
PA	(DUPO) DU PONT DE NEMOURS & CO E I.	
XX		
PI	Falco SC, Famodu OO, Thorpe CJ;	
XX		
DR	WPI: 2001-418057/44.	
XX		
DR	P-PSDB; AAE04356.	
XX		
XX	Novel nucleic acid molecules encoding aspartate kinase useful in cell	
PT	transformation and transgenic plant production procedures.	
XX		
PS	Claim 8; Page 55-56; 64pp; English.	
XX		
CC	The present invention relates to nucleic acid molecules encoding	
CC	aspartate kinase useful in cell transformation and transgenic plant	
CC	production procedures. Nucleic acid encoding aspartate kinase is useful	
CC	for producing a transgenic plant, by transforming a plant cell with a	
CC	chimeric gene comprising nucleic acid encoding aspartate kinase operably	
CC	linked to a regulatory sequence and regenerating a plant from the	
CC	transformed plant cell. The present sequence is rdrlf.pk005.f20 (FIS)	
CC	cDNA clone encoding Oryza sativa aspartate kinase	
XX		
Sequence	847 BP; 180 A; 196 C; 270 G; 201 T; 0 U; 0 Other;	

Query Match	18.7%	Score 371.2	DB 5	Length 847
Best Local Similarity	73.8%	Pred. No. 5.7e-99		
Matches 485	Conservative	0	Mismatches 169	Indels 3
			Gaps	1
Qy	274	GCTCAGCCGTGATGAAGATTGCGGGGCTCTCGGTGCTCGGCCGAGATGACTGA	333	
Db	182	GTTCAACCGTCGCCATGAAAGTTGGGGGGTCTCGGTGAGCTCGGCCGAGCGAGATGCGGA	241	
Qy	334	GGTGGCGGGGCTCATCTGACGTTCCCGAGAGACGGCCGCTGCTCTCTCCAT	393	
Db	242	GGTGGCGGATCTCATCTGACGTTCCCGAGAGACCTCCGTTGTTGTTCTCCGCCAT	301	
Qy	394	GGGGAAGACCAACCAACACTTCTCTTGCTGAGAGAGACGTAGAGGTGGAATTAT	453	
Db	302	GGGGAAGACCAACCAATTAACCTCTCTTGCCGAGAGAGAGGCTGTGAGCTGCGGCCCC	361	
Qy	454	CCATTTTCTGAAATCGAAGAGTGGAAATATGTCTAAAGCTCATATCAAGCGTGA	513	
Db	362	GAAGGCGTGTGAATTTCCCGAGCTGCAGATATTAAGAGAGCTCCATGTGACTATTA	421	
Qy	514	TGAATCTGACTTCCAGAAATCTGTATATCAAGCTTT--ATGAACGTGAGCAATT	570	
Db	422	TGAGCTTGATGTGATATATGATGATTTGTTTCAAGTTATTTGGAAGATTTGGAACAATCT	481	
Qy	571	GAAAGGTATGCTATGATGAAAGACCTGAGCGCTTAGACCAAGTACTACCTGTTTCAT	630	
Db	482	TAAAGGTATGCTATGATGAAAGACCACTACCTTAGACCAAGGATTAACCTGTTTCAT	541	
Qy	631	TGAGAGATCATGTCCACACAGATTTTTCGCTTATTTGAACAAATTCGTACAGG	690	
Db	542	CGGTGATCATGTCTCAAGAAATTTTTCGCTATTTTGAATTAACCTTGGGAAAAGG	601	
Qy	691	ACGGCAGTATGACGACTTGAATTTGTTTCATTCAACTGATGAATTTGTATGCGGA	750	
Db	602	TCGGCAGTATGATGATTTGATTTATTTGGCTTTATACACTAGATGATTTCAACATCGGA	661	
Qy	751	TATCTTAGAAGAACCTTACCTGCGTGTGCGAAGACCTCATGGGAGCTGATACAGA	810	
Db	662	CATTTCTTAGAAGCAATATCTGCTGTTCGAAAGAGGTTCACGGGTGATTTGATACGA	721	
Qy	811	TCACAGATACCTGTTGTTACTGAGTTCCTTGGAGAGGCTGGAATTCGTGCTGAAC	870	
Db	722	CCCTGTATTCCTATAGTTACTGCTTCTCTGAAAAGATGGAATTCATGTCTGTAC	781	
Qy	871	TACTTAGGCGAGGTGAGTGAATGACCTGACCTGACCAACATTTGTAAGACCTTGGG	927	
Db	782	CATTTTAGGCGAGGTGAGTGAATGACCTGACCTGACCAACATTTGTAAGACCTTGGG	838	

RESULT 13

ABX84348	
ID	ABX84348 standard, cDNA, 262 BP.
XX	
AC	ABX84348;
XX	
DT	24-APR-2003 (first entry)
XX	
DE	Corn ear-derived polynucleotide (cpd) #2608.
XX	
KM	Corn ear-derived polynucleotide; cdp, cDNA library; SATMON022; SATMON023
KM	structural gene; functional gene; regulatory gene;
KM	Corn ear-specific profile; gene transcription; gene expression;
KM	hybrid plant; desirable trait expression; plant breeding program;
KM	interference; desired characteristic; growth; development;
KM	disease resistance; environmental adaptability; quality; yield;
KM	multigene trait; plant; gene; ss.
XX	
OS	Zea mays.
XX	
FN	US6476212-B1.
XX	
PD	05-NOV-2002.

XX 14-MAY-1999; 99US-00313294.
 PF 26-MAY-1998; 98US-0086722P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Lalgudi RV, Ito LY, Sherman BK;
 XX WPI: 2003-208840/20.
 DR Novel purified corn-ear derived polynucleotide useful as hybridization
 XX probe for detecting polynucleotide in sample, and for identifying,
 PT evaluating, and altering desired characteristics associated with growth,
 PT development.
 XX Example; SEQ ID NO 2808; 390pp; English.
 PS The present invention relates to the isolation of corn ear-derived
 CC polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
 CC and SATMON023. Some of the cdfs uniquely identify structural, functional,
 CC and regulatory genes of corn ear. The polynucleotide sequences are
 CC useful for detecting cdfs in a sample, for producing a corn ear-specific
 CC profile of gene transcription, for detecting altered gene expression in
 CC inbred or hybrid plants, and for screening several molecules for specific
 CC binding to the polynucleotide. The cdfs are useful to identify, isolate,
 CC or extend identical or related corn-ear nucleic acid sequences from DNA
 CC libraries, and in nucleic acid amplification or hybridization techniques
 CC to follow the expression of desirable traits through plant breeding
 CC programs. Preferably, the cdfs are used to identify, evaluate, alter, or
 CC follow the inheritance of desired characteristics associated with growth
 CC and development, disease resistance, environmental adaptability, quality,
 CC and yield of corn. The cdfs are also useful as molecular markers for
 CC studying inheritance and multigene traits in a plant breeding program.
 CC The cdfs are useful for producing purified corn-ear polypeptides by
 CC recombinant techniques. They are also useful in diagnostic assays to
 CC detect or confirm conditions or diseases associated with abnormal levels
 CC of cdp expression. ABX81541-ABX89140 represent corn ear-derived
 CC polynucleotides (cpds) of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the USPTO web site at
 CC segdata.uspto.gov/psipdsIDentry.html
 CC
 XX Sequence 282 BP; 87 A; 50 C; 64 G; 81 T; 0 U; 0 Other;
 SQ
 Query Match 12.3%; Score 244.2; DB 7; Length 282;
 Best Local Similarity 97.1%; Pred. No. 1.4e-61;
 Matches 271; Conservative 0; Mismatches 3; Indels 5; Gaps 2;
 QY 449 GTTATCCATGTTCTGGAATCGAAGAGTGAATATGTCACCAAGCCTACATCAAGACG 508
 DB 1 GTTATCCATGTTCTGGAATCGAAGAGTGAATATGTCACCAAGCCTACATCAAGACG 60
 QY 509 GTGATGAACCTTGACCTTCCAGATCTGTATATCAAG---CCTTATGATGAGCA 564
 DB 61 GTGATGAACCTTGACCTTCCAG-ATCTGTATATCAAGACATGCTAATGATCAAGCA 119
 QY 565 ACTATTGAAAGTATCGCTATGATGAAGAGCTGACCGCTGAGCCAGTACATACCTGT 624
 DB 120 ACTATTGAAAGTATCGCTATGATGAAGAGCTGACCGCTGAGCCAGTACATACCTGT 179
 QY 625 TTCAATTGAGAGATGATGTCACCCAGGATTTTCTGCTATTATTGAACAAATTCGTGT 684
 DB 180 TTCAATTGAGAGATGATGTCACCCAGGATTTTCTGCTATTATTGAACAAATTCGTGT 239
 QY 685 CAAGGACGCGAGTATGAGCATTTGATGATGCTTAT 723
 DB 240 CAAGGACGCGAGTATGAGCATTTGATGATGCTTAT 278

XX ABE84393;
 XX 24-APR-2003 (first entry)
 XX Corn ear-derived polynucleotide (cpd) #2853.
 DE Corn ear-derived polynucleotide (cpd) #2853.
 XX
 KW Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
 KW structural gene; functional gene; regulatory gene;
 KW corn ear-specific profile; gene transcription; gene expression;
 KW hybrid plant; desirable trait expression; plant breeding program;
 KW inheritance; desirable characteristic; growth; development;
 KW disease resistance; environmental adaptability; quality; yield;
 KW multigene trait; plant; gene; ss.
 XX
 OS Zea mays.
 XX
 XX US6476212-B1.
 XX
 XX 05-NOV-2002.
 XX
 XX 14-MAY-1999; 99US-00313294.
 XX
 XX 26-MAY-1998; 98US-0086722P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX Lalgudi RV, Ito LY, Sherman BK;
 XX WPI: 2003-208840/20.
 DR Novel purified corn-ear derived polynucleotide useful as hybridization
 XX probe for detecting polynucleotide in sample, and for identifying,
 PT evaluating, and altering desired characteristics associated with growth,
 PT development.
 XX Example; SEQ ID NO 2853; 390pp; English.
 PS The present invention relates to the isolation of corn ear-derived
 CC polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
 CC and SATMON023. Some of the cdfs uniquely identify structural, functional,
 CC and regulatory genes of corn ear. The polynucleotide sequences are
 CC useful for detecting cdfs in a sample, for producing a corn ear-specific
 CC profile of gene transcription, for detecting altered gene expression in
 CC inbred or hybrid plants, and for screening several molecules for specific
 CC binding to the polynucleotide. The cdfs are useful to identify, isolate,
 CC or extend identical or related corn-ear nucleic acid sequences from DNA
 CC libraries, and in nucleic acid amplification or hybridization techniques
 CC to follow the expression of desirable traits through plant breeding
 CC programs. Preferably, the cdfs are used to identify, evaluate, alter, or
 CC follow the inheritance of desired characteristics associated with growth
 CC and development, disease resistance, environmental adaptability, quality,
 CC and yield of corn. The cdfs are also useful as molecular markers for
 CC studying inheritance and multigene traits in a plant breeding program.
 CC The cdfs are useful for producing purified corn-ear polypeptides by
 CC recombinant techniques. They are also useful in diagnostic assays to
 CC detect or confirm conditions or diseases associated with abnormal levels
 CC of cdp expression. ABX81541-ABX89140 represent corn ear-derived
 CC polynucleotides (cpds) of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the USPTO web site at
 CC segdata.uspto.gov/psipdsIDentry.html
 CC
 XX Sequence 270 BP; 83 A; 48 C; 63 G; 76 T; 0 U; 0 Other;
 SQ
 Query Match 11.3%; Score 225.2; DB 7; Length 270;
 Best Local Similarity 96.7%; Pred. No. 5.7e-56;
 Matches 263; Conservative 0; Mismatches 3; Indels 6; Gaps 3;
 QY 449 GTTATCCATGTTCTGGAATCGAAGAGTGAATATGTCACCAAGCCTACATCAAGACG 508
 DB 1 GTTATCCATGTTCTGGAATCGAAGAGTGAATATGTCACCAAGCCTACATCAAGACG 60

QY 509 GTGATGAACCTGGACTCTCCAAANATCTGTAATACAG---CCTTATGAACCTGAGCA 564
 Db 61 GTGATGAACCTGGACTCTCCAG-ATCTGTAAATACAGACATGCTAGATGAACCTGAGCA 119
 QY 565 ACTTGAAGGATATGCTCTGATGTAAGAGCTGACGCTAGGACCACTGACTTGT 624
 Db 120 ACTATTAAGATATGCTCTGATGTAAGAGCTGACGCTAGGACCACTGACTTGT 179
 QY 625 TTCAATTTGGAATGATGTCACACAGATTTTTCCTTATTTGAACAAATTCGTGT 684
 Db 180 TTCAATTTGGAATGATGTCACACAGATTTTTCCTTATTTGAACAAATTCGTGT 238
 QY 685 CAAGGACGGGAGATGATGAGCATTTGATTTG 716
 Db 239 CAAGGACGGGAGATGATGAGCATTTGATTTG 270

RESULT 15

AD08623/c
 ID AD08623 standard; cDNA; 455 BP.

AC AD08623;

DT 04-SEP-2001 (first entry)

DE Oryza sativa aspartate kinase encoding rdlf.pk005.f20 (EST) cDNA clone.

XX Rice; aspartate kinase; cell transformation; transgenic plant;

KM rdlf.pk005.f20 clone; ss.

OS Oryza sativa.

XX Key Location/Qualifiers
 FT complement (1.342)
 FT /tag= a
 FT /product= "Oryza sativa rdlf.pk005.f20 (EST) aspartate kinase"

XX WO200146393-A2.

XX 28-JUN-2001.

XX 19-DEC-2000; 2000WO-US034396.

XX 21-DEC-1999; 99US-0172944P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Falco SC, Famodu OO, Thorpe CJ;

XX WPI; 2001-418057/44.

XX P-PSDB; AAE04355.

XX Novel nucleic acid molecules encoding aspartate kinase useful in cell transformation and transgenic plant production procedures.

XX Claim 8; Page 55; 64pp; English.

XX The present invention relates to nucleic acid molecules encoding aspartate kinase useful in cell transformation and transgenic plant production procedures. Nucleic acid encoding aspartate kinase is useful for producing a transgenic plant, by transforming a plant cell with a chimeric gene comprising nucleic acid encoding aspartate kinase operably linked to a regulatory sequence and regenerating a plant from the transformed plant cell. The present sequence is rdlf.pk005.f20 (EST) cDNA clone encoding Oryza sativa aspartate kinase

XX Sequence 455 BP; 143 A; 112 C; 74 G; 119 T; 0 U; 7 Other;

Query Match 10.3%; Score 204.2; DB 5; Length 455;

Best Local Similarity 78.8%; Pred. No. 1.3e-49; Matches 256; Conservative 0; Mismatches 68; Indels 1; Gaps 1;

QY 602 CTTAGACCAAGTACTACTTGTTCATTGGAAGATGATGCCACAGATTTTTC- 660
 Db 325 CTTAGACACGGGATTTACTTGTTCCTGCTGATGATGATCTACAGAAATTTGCT 266
 QY 661 TGCTATTGAAACAAATTCGTGTCAGGACGCGAGATGACGATTTGATTTGTT 720
 Db 265 TGATATTGAAATPAACTTGGAAAAAGCTGGAGATGATGATTTGATTTGGCT 206
 QY 721 CATTACACTGATGATTTGTAATGCGGATATCTTAAGCAACTATCTGCTTTC 780
 Db 205 TATTACTACTGATGATTTTACAAATGCGGACATTTCTTAAGCAATATCTGCTTTC 146
 QY 781 GAAGAGACTTCATGGGGACTGATACAGGATCAGACGATACCTGTTTACTGGGTTCT 840
 Db 145 AAGAGGTTACAGGATGATGATTTGACGACCTGCTATTTCTAATGATGATGTTTCT 86
 QY 841 TGGAAAGGCTGGAATCTGTGCTGTAATCTTTAGGCGGAGTGTATGACTTGAC 900
 Db 85 TGGAAAAAGATGGAATTCATGTGCTGTCAACCACTTTAGCAGAGGTGACGACTTGAC 26
 QY 901 TGCTACACCATTTGATPAAAGCCTTG 925
 Db 25 TGCTACACCATTTGCAAGCGTTG 1

Search completed: March 23, 2004, 07:25:12
 Job time : 803 secs

ESUUI 2
S-09-313-294A-2853

Sequence 2853, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Ialjudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley X.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 2853
LENGTH: 270
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6476212 700553382H1
US-09-313-294A-2853

Query Match 11.3%; Score 225.2; DB 4; Length 270;
Best Local Similarity 96.7%; Pred. No. 4.2e-57;
Matches 263; Conservative 0; Mismatches 3; Indels 6; Gaps 3;

QY 449 GTATCCATGTTCTGTAATCGAAGTGGATATGTCAAAAGCCTACATATCAAGACG 508
DB 1 GTTTCATGTTCTGTAATCGAAGTGGATATGTCAAAAGCCTACATATCAAGACG 60
QY 509 GTGATGACCTTGACTTCCAAAGATCTGTATACAG---CCTTATGAACGTGACGA 564
DB 61 GTGATGACCTTGACTTCCAAAG-ACCTGTATACAAAGCATGATGATGATGACGA 119
QY 565 ACTATGAAAGGTATGCGCTATGATGAAGAGCTGACGCTAGACCAAGTACTACCTGT 624
DB 120 ACTATGAAAGGTATGCGCTATGATGAAGAGCTGACGCTAGACCAAGTACTACCTGT 179
QY 625 TTCAATTGAGAAATGATGTCACCAAGATTTTCTGCTTATTTGAACAAATTCGTGT 684
DB 180 TTCAATTGAGAAATGATGTCACCAAGATTTTCTGCTTATTTGAACAAATTCGTGT 238
QY 685 CAAGGACGGACGATGACGCAATTTGATTTG 716
DB 239 CAAGGACGGACGATGACGCAATTTGATTTG 270

RESULT 3
US-08-916-421B-1/c
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
APPLICANT: Bull et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
Patent No. 6503729
FILE REFERENCE: PB275
CURRENT APPLICATION NUMBER: US/08/916,421B
CURRENT FILING DATE: 1997-08-22
PRIORITY APPLICATION NUMBER: US 60/024,428
PRIORITY FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc_feature
LOCATION: (28222)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc_feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g

```

LOCATION: (600992) .. (600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (622708) .. (622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657081) .. (657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657203) .. (657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (674435) .. (674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (682442) .. (682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (713652) .. (713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (741684) .. (741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779455) .. (779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779676) .. (779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (855539) .. (855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (871619) .. (871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1084830) .. (1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1096846) .. (1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1119881) .. (1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1130881) .. (1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1310988) .. (1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224) .. (1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349473) .. (1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349491) .. (1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1470091) .. (1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1569020) .. (1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1602912) .. (1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1603734) .. (1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1637998) .. (1637998)

```

```

OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1664854) .. (1664855)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

```

```

Query Match      6.5%  Score 129.8; DB 4; Length 1664976;
Best Local Similarity 49.2%; Pred. No. 2e-26;
Matches 440; Conservative 0; Mismatches 434; Indels 21; Gaps 3;

```

```

QY 389 GCCATGGGAAACCCACCAACACCTTCTCCTGCTGGAAGAAAGCGAGTGGTGA 448
DB 963821 GCATTGGTGGAGATATCTCAACCAAGCTTGAAGATATCGCAAGAGAGAT 963762
QY 449 GTTATCCATGTTTTCGAAATCGAAGAGTGAATATGCTCAAAAGCCTACATCAAGACG 508
DB 963761 TTTATTAATTTATTTAGAGAAACACTCAAAAGCTTAAAGAGCTATMAAATCAGAA 963702
QY 509 GTGATGAACCTGACCTTCCAGNACTGTAAATCAAGCCTTATGAACTGAGCACTA 568
DB 963701 GAATTAAGAAGAGTAAATAATTAATGACAGAGATTGAGATTAAGAAAGTT 963642
QY 569 TTGAAGGTATTCCTATATATAAGAGCTGACGCTTGAACGCTGACTCTGTTTCA 628
DB 963641 TTAAATGGTGTAGCATCTTAGGGAGCTTACCAAGTCAAGAGCTATATATATCA 963582
QY 629 TTGAGAAATGCATCTCCACAGATTTTTCGCTATTTGAACAAAATTCGTGCAAG 688
DB 963581 TTGAGAGAGAGTCTCTCTCAACATATTAAGTGAAGCTATTAAGATTAAGAGAAAG 963522
QY 689 GCACGACATATGACGCAATTTGATATGTTTCACTTCAACTGATGATTTGGTAATCG 748
DB 963521 TCTATTTGCTTGAAGAGGAGAAAGCAGAAATTAACGATTAACACTTTGGAAGTCA 963462
QY 749 GATATCTTAGAAGCAACTATCTGCTGTCGAAGAGACTTCAAGGAGCTGAGATACAG 808
DB 963461 AGAGTTAAAGATTTAGAGTTTAAAGAGATTTTATCA-----TTATTAATA 963414
QY 809 GATCCAGGATACCTGTGTTTACTGCGGTTCTTGGAAAGGCTGGAATCTGTGCTGTA 868
DB 963413 GAGGCAATTAATCCAGTGTGTACAGATTTTATAGAA-----CCACTGAAGGGGTATATA 963357
QY 869 ACTACTTATAGCCGAGGAGTGTAGTGAAGTGTGCTGCTCAACCATTTGTAAAGCCTTGGGA 928
DB 963356 ACACCTTGAAGAGGCGGAAGTGAATTAAGTCAAGCCTTTAATTTGTTATGCTTGAT 963297
QY 929 CTGAGGAATTCAGATATGAAAGATGTTGATGTTGATCTTACTTGTGATCCAAATATC 988
DB 963296 GCAGATATATTAATTTGAGCAGATGTTTCTGAGTTTATACAGAGCCCAAGATTA 963237
QY 989 TACCCACATGCAAAACACTGTTCCATCTTAATTGAAAGGCGACAGAACTTGCTTAT 1048
DB 963236 GTTCTTACTGCAAGAGAAATTCAAAAGTTGTTACATAGAGGCTATGAAATTTGCAATC 963177
QY 1049 TTTGTGCTCAGGTTTTCATATCCATGCAATGATGAGACCTGTAGAGAGGTATATCCA 1108
DB 963176 TTTGAGGTCTAAGTTTTCATATCCAAAGCTATAGAGCAGTATGAGAAAGGCAATCCA 963117
QY 1109 GTTAGGGTTAAGATTTATCAACCTTAAGCTCCAGGACCCCTTATTAACGACAAAGA 1168
DB 963116 ATATTGTAAAGAAATTAATTTAGGCAAGAGTGAAGAAATTTAATTAACCAAGATATG 963057
QY 1169 GACATGATTAANGTCTGGTTGTACTAATGATGCAATGCTCAAGTCAATGCTCATATG 1228
DB 963056 GAGATGAGTAGATGATAGT-----AAAGCTATCAACAATAAATAATGCTGCTTA 963003
QY 1229 TTGAGCATTTGAGCACTGCGATGCTTGTGATGATGTTTTCGCGCAAGGCTAT 1283
DB 963002 ATAAACATTTTGGAGCTGGAGTGTGTGAGTTAGTGAACAGCGGCAAGGATAT 962948

```

```

RESULT 4
US-09-790-988-1/c

```

```

; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGEMOBU, SHUJI
; APPLICANT: MATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHITUKU
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

```

```

Query Match      4.3%; Score 86; DB 4; Length 640681;
Best Local Similarity 48.5%; Pred. No. 1.2e-13;
Matches 236; Conservative 0; Mismatches 251; Indels 0; Gaps 0;

```

```

QY 699 ATGACGCAATTGATTTGGTTTCATTTACACTGATGAAATTTGGTAATGCGATATCTTAG 758
DB 211266 ATGTACTATTATTATCTCTGTTAAAAATTAGTTGCTATAGTGATATTTCTTGATT 211207
QY 759 AAGCAACCTATCTCTGTTGGGAGAGACCTTCATGGGAGCTGATACAGATTCAGCGA 818
DB 211206 CAACGTGATATATCTAATCTAAAAAATATTCAGACATATAATTAATCAAGTA 211147
QY 819 TACCTGTGTTACTGAGGTTCCCTGGGAGAGGCTGGAATCTGTCGTACTACTTAG 878
DB 211146 ATATTATTATTAATGCTGCTGTTTATTGCGGAAATTAAGATAAAAATAGTGTATTAG 211087
QY 879 GCCGAGGTGTAGTACTGACTGCTGCTACACCAATGTGTAAGCCCTGGGACTGAGAGAA 938
DB 211086 GACGAAATGTTCTGATATTACGCTGCGGTAGCGGCTGTTGATGCGAATGTT 211027
QY 939 TTCAGGTATGGAAGAATGTGATGCTGTAATCTTGTGATCCAAATATCAOCCCATG 998
DB 211026 GTGAATCTGACTGATGTCGAGGAGTTTACTTCTGATCCAGAAAGTTCCAAATG 210967
QY 999 CAAGACTGTTCCACTACTTAACATTTGAGAGGCGACAACTTGCTTATTGGTCTC 1058
DB 210966 CTCGTTATTTAAATCAATATCATATCAAGAGCATGSACTGCTTATTGCGTCTA 210907
QY 1059 AGGTTTGCATCCAAATCGATGAGACTGCTAGAGAAAGGATATCCAGTTAGGTTA 1118
DB 210906 AGGTGCTACACCTGCTGATATGTAACCAATGCTCAATTTAAATTCATGTTTAATTA 210847
QY 1119 AGAATTCATACACCTTAAGCTCCAGGCACTTATTCCAGCAAAAGAGACATGATA 1178
DB 210846 AAAATACAAATATATTTAAATCTATTGGAACATTAATTTGGAACAAAGCTTCCAAA 210787
QY 1179 ANGCTCT 1185
DB 210786 AAGATTT 210780

```

```

RESULT 5
US-09-543-681A-3076
; Sequence 3076, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05

```

```

; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 3076
; LENGTH: 1383
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-3076

```

```

Query Match      4.1%; Score 81.6; DB 4; Length 1383;
Best Local Similarity 54.3%; Pred. No. 6.3e-14;
Matches 165; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

```

```

QY 869 ACTACTTTAGGCGGAGGTGTAGTGTGATGCTGATACCAACCACTTGTAAAGCTTGGGA 928
DB 616 ACCACTTAGGGCTGCGGAGATGATTACTGCGGCACTGATTTGTTAGAGCATTTAGGG 675
QY 929 CTGAGAGAAATTCAGTATGAGAAAGATGTGATGTTACTTACTTGATCCAAATATC 988
DB 676 ATGTACGTCATATTGAGACTGATGTCTGTATTATTAAGCACTGATCCAGAAATA 735
QY 989 TACCCATATGAAAGATGTTCCACTTACTTAACTTTGAAGAGCCACAGAACTTGCTAT 1048
DB 736 GTCCCGAAGCAATCGTATTTGATCATATGCTTTTGAAGAGCGGCTAAATGCAAC 795
QY 1049 TTTGATGCTCAGTTTTCATCCATCCATGATGAGACTGCTAGAGAGGTGATATTCCA 1108
DB 796 TTGATGCTAAATTTCCATCCCGCAACTTATTACCTGCAAGTGGTATGCGATCCCT 855
QY 1109 GTTAGGTTAAGATTATATCAACCTTAAGCTCCAGGCACTTATTCCAGACAAAGA 1168
DB 856 GTGTTTGGCTCCTACAGCAAGACAGAGAGGTGACATGTTGTGACAAACT 915
QY 1169 GACA 1172
DB 916 GATA 919

```

```

RESULT 6
US-09-543-681A-2143
; Sequence 2143, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 2143
; LENGTH: 2589
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-2143

```

```

Query Match      4.0%; Score 80; DB 4; Length 2589;
Best Local Similarity 60.6%; Pred. No. 2.8e-13;
Matches 131; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

```

```

QY 940 TCAGTATGAGAAATGTTGATGTTACTTGTGATCCAAATATCTACCCATGTC 999
DB 738 TGAATTTGAGCGATTTGATGTTGTTTACCTGTGATCCAGGCTGTCAGATGC 857
QY 1000 AAAGCTGTTCCATCTTAACATTTGAAGAGCCACAGAACTTGCTTATTGGTCTCA 1059
DB 858 ACACCTATTAAGAGCATGTCCTATCAAGAGGAGGAGGAACTTCTTACTTTGGTCCAA 917
QY 1060 GATTGTGATCCAAATCGATGAGACTGCTAGAGAGGTGATATCCAGTTAGGGTTAA 1119
DB 918 AGTCTTCATCCAGCACTATGCACTATTTGCTCAATTTCAATATCCCTTTTAATTA 977

```

QY 1120 GAATCATACAACCTTAAGCTCCAGGACCCCTTAT 1155
DB 978 AATTACGGGCAATCCAGAGCACCGGTACGTTAT 1013

RESULT 7

US-07-746-705A-16
Sequence 16, Application US/07746705A
Patent No. 5451516
GENERAL INFORMATION:
APPLICANT: Matthews, Benjamin F.
APPLICANT: Weisemann, Jane M.
TITLE OF INVENTION: A Recombinant DNA Molecule Encoding
TITLE OF INVENTION: A Bifunctional Plant Enzyme: Aspartokinase and Homoserine
TITLE OF INVENTION: Dehydrogenase
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Janelle S. Graeter
STREET: Bldg. 005, Room 402, BARC-W
CITY: Beltsville
STATE: Maryland
COUNTRY: USA
ZIP: 20705
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/746,705A
FILING DATE: 19910816
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Graeter, Janelle S.
REGISTRATION NUMBER: 35,024
REFERENCE/DOCKET NUMBER: 4000.91
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301)504-5676
TELEFAX: (301)504-5060
INFORMATION FOR SEQ. ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 2915 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 2..2593
US-07-746-705A-16

Query Match 3.9%; Score 77.8; DB 1; Length 2915;
Best Local Similarity 53.9%; Pred. No. 1.3e-12;
Matches 160; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 869 ACTACTTTAGGCCGAGGTGTAGTACTGCTACCAACCATTTGTAAGCCTTGGA 928
DB 695 ACACTTTGAAAAGAGAGAGAGTACTTTCTGCCCGCTATATGGTCTTATTAAG 754
QY 929 CTGAGGAATTCAGTATGAGAAAGTGTGATGTGTAATCTTACTTGATGCCAATATC 988
DB 755 GCTGTCAAGTCAAGATTGAGCTGATGTATATGTGTATATAGTCAGATCTCGAAAA 814
QY 989 TACCACATGCAAGAAGCTGTCCATCTTAACATTTGAAGAGCCACAGAACTTGCTAT 1048
DB 815 GTTAGTAGGCTGTGTATTAAGACATTTGTTATCAAGAAAGCTGGAGATGTCAAT 874
QY 1049 TTGTGTCTCAGGTTTGTGATCCACATGATGAGAAAGCTGCTAGAGAAAGTGATATCCA 1108
DB 875 TTGGGGCTTAATGTGTATCATCCCGTACTATCATTCCTGTGTAGTCGATATGACATTTCCA 934

QY 1109 GTTAGGTTAGAAATTCATACACCTTAAGCTCCAGGACCCCTTATTACAGCAA 1165
DB 935 ATTGTATATAGAAATATATATCATCACTATCTCTCCGGAGCAATGATATGCGAGAA 991

RESULT 8

US-08-380-182-18
Sequence 18, Application US/08380182
Patent No. 5858749
GENERAL INFORMATION:
APPLICANT: Matthews, Benjamin F.
APPLICANT: Weisemann, Jane M.
TITLE OF INVENTION: A Bifunctional Protein From Carrots
TITLE OF INVENTION: (Daucus carota) with Aspartokinase and Homoserine
TITLE OF INVENTION: Dehydrogenase Activities
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Janelle S. Graeter
STREET: Room 411, Bldg. 005, BARC-W
CITY: Beltsville
STATE: Maryland
COUNTRY: USA
ZIP: 20705
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/380,182
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Graeter, Janelle S.
REGISTRATION NUMBER: 35,024
REFERENCE/DOCKET NUMBER: 0226.94
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-504-6629
TELEFAX: 301-504-5060
INFORMATION FOR SEQ. ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2915 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: Daucus carota
FEATURE:
NAME/KEY: CDS
LOCATION: 2..2593
US-08-380-182-18

Query Match 3.9%; Score 77.8; DB 2; Length 2915;
Best Local Similarity 53.9%; Pred. No. 1.3e-12;
Matches 160; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 869 ACTACTTTAGGCCGAGGTGTAGTACTGCTCTCAACCATTTGTAAGCCTTGGA 928
DB 695 ACACTTTGAAAAGAGAGAGAGTACTTTCTGCCCGCTATATGGTCTTATTAAG 754
QY 929 CTGAGGAATTCAGTATGAGAAAGTGTGATGTGTAATCTTACTTGATGCCAATATC 988
DB 755 GCTGTCAAGTCAAGATTGAGCTGATGTATATGTGTATATAGTCAGATCTCGAAAA 814
QY 989 TACCACATGCAAGAAGCTGTCCATCTTAACATTTGAAGAGCCACAGAACTTGCTAT 1048
DB 815 GTTAGTAGGCTGTGTATTAAGACATTTGTTATCAAGAAAGCTGGAGATGTCAAT 874
QY 1049 TTGTGTCTCAGGTTTGTGATCCACATGATGAGAAAGCTGCTAGAGAAAGTGATATTTCCA 1108

DB 875 TTGGGGGCTAATGTTGTTACATCCCGTACTATCATCTCTGTGATGCCATATGATGATTCGA 934
QY 1109 GTTAGGGTTAGATTCATACACCCCTTAAGCTCCAGGACCCCTTATTACAGACAA 1165
DB 935 ATTGATATAGAAATATATATTCACCTATCTGTCGGGAACAATGATATGCCGAGAA 991

RESULT 9
US-09-221-017B-1020
Sequence 1020, Application US/09221017B
Patent No. 6444799
GENERAL INFORMATION:
APPLICANT: Ross, Bruce C.
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
NUMBER OF SEQUENCES: 1120
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FORSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P22911
FILING DATE: 09-APR-1998
APPLICATION DATA:
PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: MONTROY, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 1020:
SEQUENCE CHARACTERISTICS:
LENGTH: 3412 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORYPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc feature
LOCATION: 1...3412
US-09-221-017B-1020

Query Match 3.5%; Score 69.8; DB 4; Length 3412;
Best Local Similarity 52.2%; Pred. No. 3.5e-10;
Matches 155; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

881 CGAGTGTGAGTACTGCTACACCATTTGTAAGCCTTGGAGTGGAGAAATT 940

DB 2797 CGTGGCGGCACGACACTACAGGCGCTCTGTGATCGAGACAGACACTGAATGCTGATGAATT 2856
QY 941 CAGGATATGGAAGATGTTGATGTGTACTTGTGATTCATTAATATCAACCAATGCA 1000
DB 2857 CAGATATGAGAGATATCGAGGATGACACAAATGATCTGTGCTGAGCGTTACG 2916
QY 1001 AAGACTGTTCATCTTAATCAATTTGAAGAGCCACAGAACTTCTATTTTGTGTCTAG 1060
DB 2917 GCACCCGTTGGCGGCTCTCCATTTTGAAGAAAGCTGCCGAACCTGCACACTTCGAGCCAAA 2976
QY 1061 GTTTGATCCACATGATAGACCTGCTAGAGAAAGTATATTCAGTTAGAGGTTAAG 1120
DB 2977 ATCCCTCATCCGACCTGTATCAACCGCCGCAAGAGAAATATTCGGTACGTTTACTC 3036
QY 1121 AATCATACACCCCTTAAGCTCCAGGACCCCTTATTACAGACAAAGACATGAT 1177
DB 3037 AATGATGATATCCGATGACACCCGGGACGCTGATTTCACTGATACCGAAAAGAT 3093

RESULT 10
US-09-107-532A-2436
Sequence 2436, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7110
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 2436:
SEQUENCE CHARACTERISTICS:
LENGTH: 1437 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...1437
SEQUENCE DESCRIPTION: SEQ ID NO: 2436:
US-09-107-532A-2436

Query Match 3.5%; Score 68.8; DB 4; Length 1437;
Best Local Similarity 52.6%; Pred. No. 4.2e-10;
Matches 113; Conservative 0; Mismatches 153; Indels 3; Gaps 1;

QY 875 TTAGGCGAGGTGGTGTGATGACTGTACACCATTTGTAAGCCTGGAGCTGAGA 934
DB 628 TTGGGCGCGTGGGGAAGCAGCATTAACCGCGGATACGCGCAATGTTCCAGCGGA 687
QY 935 GAAATTCAGTATGAGAAAGTGTGATGTTACTTGTGATCCAAATATCTACCCA 994
DB 688 AGTTATGAGATTTGACGATGTTTCAGGAATCATGATGCTGATCCCGAATCATCAT 747
QY 995 CATGCAAGCTGTTCATACATTAACATTTGAAGGCGACAGACTGCTTAATTTGGT 1054
DB 748 CATCTTAATATCATGAGGAATCTTTTAAGAACTGAGAGATGCTTAATATGAGG 807
QY 1055 GCTCAGGTTTGCATCCAGATGATGAGACCTGCTAGAGAGTGAATATTCAGTTAG 1114
DB 808 GCTACGTTCTTCATAGACAGCTGTTTCCAGTTCAAGAGGCTGATATCCCATCCA 867
QY 1115 GTTAAATTCATACACCTTAAGCTCCAGGACCTTATTAACAGAC---AAGAGAC 1171
DB 868 ATTAATAATTCATATATCTTTCCTCAGAGGACAAAATATCGAATCAGAAATAGAA 927
QY 1172 ATGATATANGTCTGTTGTACTACTAG 1200
DB 928 AAGAAATGACTGACAGAAATAGCTG 956

RESULT 11
US-09-557-884-1/c
Sequence 1, Application US/09557884
Patent No. 6506581

GENERAL INFORMATION:

APPLICANT: Fleischmann et al.

TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884

FILING DATE: 25-Apr-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102

FILING DATE: JUN-5-1995

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB186P3

TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504

TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs

TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-557-884-1

Query Match 3.4%; Score 66.6; DB 4; Length 1830121;
Best Local Similarity 52.3%; Pred. No. 1.4e-07;
Matches 147; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 876 TTAGGCGAGGTGGTGTGATGACTGTACACCATTTGTAAGCCTGGAGCTGAG 935
DB 95676 TCGGTGTATATGCTCAGACATATCTGCGGCTGTTAGCCGCTGTTTAGCCCAAGTG 95617
QY 936 AATTCAGTATGAGAAAGTGTGATGTTACTTGTGATCCAAATATCTACCCAC 995
DB 95616 TATGCGAAATTTGACCCATGTGACGCGCTTTTACTTGGATCCGCGTTAGTGCCTG 95557
QY 996 ATGCAAGCTGTTCATACATTAACATTTGAAGGCGACAGACTGCTTAATTTGGTG 1055
DB 95556 ATGCAAGCTTATTAACCAACGCTTTTATCGTGAAGCATGGAACCTCTTAATTCGATG 95497
QY 1056 CTCAGGTTTGCATCCAGATGATGAGACCTGCTAGAGAGTGAATATTCAGTTAGGG 1115
DB 95496 CGAAAGTATTCATCTCTGATGATGAGCCATGTTGCCAATAATATCTTGTGTCA 95437
QY 1116 TTAAGATTCATACACCTTAAGCTCCAGGACCTTAAT 1156
DB 95436 TTAATAATTCAGAAATCTTCTGACCAAGGTTGATTAAT 95396

RESULT 12
US-09-643-990A-1/c
Sequence 1, Application US/09643990A
Patent No. 6528289

GENERAL INFORMATION:

APPLICANT: Robert D. Fleischmann

Mark D. Adams

Owen White

Hamilton O. Smith

J. Craig Venter

TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A

FILING DATE: 23-Aug-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429

FILING DATE: 1995-06-07

APPLICATION NUMBER: 08/426,787

FILING DATE: 1995-04-21

ATTORNEY/AGENT INFORMATION:

NAME: Kenley K. Hoover

REGISTRATION NUMBER: 40,302

REFERENCE/DOCKET NUMBER: PB186P1C1

TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790

TELEFAX: 310-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs

TYPE: nucleic acid
STRANDEDNESS: double

US-09-643-990A-1

TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match 3.4%; Score 66.6; DB 4; Length 1830121;
Best Local Similarity 52.3%; Pred. No. 1.4e-07;
Matches 147; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 876 TAGGCCGAGGTGTAGTACTTGACTGCTACCAACCATTTGTAAGCCTTGGGACTGAG 935
DB 95676 TCGGCTGTAATGGCTCAACACTATTCTGCGCTCTTTAGCCGCTTTTGGCGCAAGTG 95617
QY 936 AAATTCAGGTATGGAAGATGTGTAGTGTACTTACTTGTATCCAAATATCTACCCAC 995
DB 95616 TATGCGAATTTGGACGATGTGGACGGCGTTTATCTTGGATCCGCGTTTGTGCTTG 95557
QY 996 ATGCAAGACTGTCTCACTTACTTAACTTTGAAGGCGCACAGAACTTCTTATTTGGTG 1055
DB 95556 ATGACGCTTTATTAACCAACGCTTTCTTATCTGTGAAGCGATGGAACCTCTTATTTGGTG 95497
QY 1056 CTCAGGTTTGCATCCACATCGATGATGAGACCTGCTAGAGAAGGTATATCCAGTTAGGG 1115
DB 95436 CGAAGTATTCATCTCTGTGAGATGGGCCATTGTTGCCAATAATATCTTGTGTCA 95437
QY 1116 TTAGAATTCATACACCTTAAGCTCCAGGACCTTATT 1156
DB 95436 TTTAAATAATACAGAAATCTTCTGCACACGATTTGATT 95396

RESULT 13
US-08-532-828B-11
Sequence 11, Application US/08532828B

Patent No. 5688671
GENERAL INFORMATION:
APPLICANT: SUGIMOTO, Masakazu
APPLICANT: OGAWA, Yuki
APPLICANT: SUZUKI, Tomoko
APPLICANT: TANAKA, Akiko
APPLICANT: MATSUI, Hiroshi
TITLE OF INVENTION: MUTANT ASPARTOKINASE GENE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-DOS TEXT EDITOR
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532, 828B
FILING DATE: 27-OCT-1995
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: JP 5-101450
FILING DATE: 27-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: NORMAN F. OBLON
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-764-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1263 nucleotides
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Corynebacterium glutamicum
STRAIN: ATCC13869
US-08-532-828B-11

Query Match 3.2%; Score 64.2; DB 1; Length 1263;
Best Local Similarity 49.1%; Pred. No. 9e-09;
Matches 199; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

QY 757 AGAAGCAACCTATCTGCTGTGCGAAGACTTCATGCGGACTGATACAGATCCAGC 816
DB 327 AAACGACGATTTGTAAGCTCACACCGGCTGCTGCGTAAGCACTCGATGAGGCAA 386
QY 817 GATACCTGTGTACTAGGCGTCTCTTGGGAAGGCTGGAATCTGTGCTTACTACTTT 876
DB 387 GATCTGATTTGCTGCTGTTTCAAGGTGTTAATAAGAAACCCGATGTACCAAGTT 446
QY 877 AGCGGAGGTGTAGTACTGACTGCTCAACCACTTGTAAAGCCTTGGACTGAGA 936
DB 447 GGTGCTGTGTGTTCTGACACCACTGACGTGGGTGGCAGCTGCTTGAAGCTGATGT 506
QY 937 AATTCAGTATGGAAGATGTGTATGTGTACTTATTTGTATCCAAATATCTACCCACA 996
DB 507 GTGTGATTTACTCTGACGCTTGACGCTGTGTATACCGCTGACCGCGCATGTTCTTA 566
QY 997 TGCAAGACTGTCTCACTTACTTAACTTTGAAGGCGCACAGAACTTGTCTTATTTGGTG 1056
DB 567 TGCAAGAACTGGAAGAACTGACCTTGAAGAAATGCTGAACTTGTGCTGTTGGCTC 626
QY 1057 TCAGTTTTCATCCACATCGATGAGACCTGCTAGAGAAGGTATATCCAGTTAGGT 1116
DB 627 CAAGATTTGTGTGTGCGAGGTGTAAGTCTGCTGATTCATATGTCACCTTGGCGT 686
QY 1117 TAAGATTCATACACCTTAAGCTCCAGGACCTTATTACAG 1161
DB 687 ACGCTGCTTATTA---GTAAATATCCGCGCACTTGAATTGCGG 728

RESULT 14

US-08-532-828B-12
Sequence 12, Application US/08532828B
Patent No. 5688671

GENERAL INFORMATION:
APPLICANT: SUGIMOTO, Masakazu
APPLICANT: OGAWA, Yuki
APPLICANT: SUZUKI, Tomoko
APPLICANT: TANAKA, Akiko
APPLICANT: MATSUI, Hiroshi
TITLE OF INVENTION: MUTANT ASPARTOKINASE GENE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-DOS TEXT EDITOR
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532, 828B
FILING DATE: 27-OCT-1995
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: JP 5-101450
FILING DATE: 27-APR-1993
ATTORNEY/AGENT INFORMATION:

Wed Mar 24 14:12:53 2004

us-09-890-813-5.rnt

Page 9

NAME: NORMAN F. OBLON
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-764-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1263 nucleotides
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Corynebacterium glutamicum
STRAIN: AJ3463
US-08-532-828B-12

Query Match
Best Local Similarity 49.1%; Score 64.2; DB 1; Length 1263;
Matches 199; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

QY 757 AGAAGCAACTATCTCTGTTGCGAAGAGACTTCATGGGACTGATACAGATCCAGC 816
DB 327 AAACGACGACATTTGACGTACACCGGGTGTGTGCGTGAACACATCGATGAGGCA 386
QY 817 GATACCTGTTGTTACTGGTTCCTTGCGAAGGCGTGAATCTGTGTGTAATCTTT 876
DB 367 GATCTGCATTTGTTGCTGTTTTCAGGGTGTATTAAGAAACCCGCGATGACACAGTT 446
QY 877 AGCGCGAGGTGTGATGACTGACTGTACAAACCATTTGTAAGCTTTGGAGTGAAGA 936
DB 447 GGGTGTGTGTGTTGTCGACACACGATTCGCTTGACGCTGTTGAACGCTGATGT 506
QY 937 AATTCAGTATGGAAGATGTGATGATGATGATGATGATGATGATGATGATGATGAT 996
DB 507 GTGTGATTTACTGCGAGCTGACGCTGTGTATACCCCTGACCCGCGCATCTCTTAA 566
QY 997 TCGAAGACTGTTCATCTTAACTTAACATTTGAAGAGCCACAGAACTTGTATTTTGTG 1056
DB 567 TCGACGAGAGCTGGAAGAAAGCTCAGCTTGAAGAAATGCTGGAATCTGTGTGTGCTC 626
QY 1057 TCAGGTTTGCATCCCAATGATGAGACCTGTGAGAAAGTGATATTCAGTTAGGT 1116
DB 627 CAAGATTTTGTGCTGCGCAGATGTAATACGCTGTGATCAATGTGCGCACTTGCCT 686
QY 1117 TAGAATTCATCAACCTTAAGCTCCAGGCAACCTTATTACAG 1161
DB 687 AGCGTGTCTTATA---GTAATGATCCCGGCACTTTGATTTGCCG 728

RESULT 15
US-08-532-828B-1
Sequence 1, Application US/08512828B
Patent No. 5686671
GENERAL INFORMATION:
APPLICANT: SUGIMOTO, Masakazu
APPLICANT: OGAWA, Yutai
APPLICANT: SUZUKI, Tomoko
APPLICANT: TANAKA, Akiko
APPLICANT: MATSUI, Hiroshi
TITLE OF INVENTION: MUTANT ASPARTOKINASE GENE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT, P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-DOS TEXT EDITOR
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532,828B
FILING DATE: 27-OCT-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-101450
FILING DATE: 27-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: NORMAN F. OBLON
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-764-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1643 nucleotides
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Corynebacterium glutamicum
STRAIN: ATCC 13639
US-08-532-828B-1

Query Match
Best Local Similarity 49.1%; Score 64.2; DB 1; Length 1643;
Matches 199; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

QY 757 AGAAGCAACTATCTCTGTTGCGAAGAGACTTCATGGGACTGATACAGATCCAGC 816
DB 543 AAACGACGACATTTGACGTACACCGGGTGTGTGCGTGAACACATCGATGAGGCA 602
QY 817 GATACCTGTTGTTACTGGTTCCTTGCGAAGGCGTGAATCTGTGTGTAATCTTT 876
DB 603 GATCTGCATTTGTTGCTGTTTTCAGGGTGTATTAAGAAACCCGCGATGACACAGTT 662
QY 877 AGCGCGAGGTGTGATGACTGACTGTACAAACCATTTGTAAGCTTTGGAGTGAAGA 936
DB 663 GGGTGTGTGTGTTGTCGACACACGATTCGCTTGACGCTGTTGAACGCTGATGT 722
QY 937 AATTCAGTATGGAAGATGTGATGATGATGATGATGATGATGATGATGATGATGAT 996
DB 723 GTGTGATTTACTGCGAGCTGTGACGCTGTGTATACCGCTGACCCGCGCATCTCTTAA 782
QY 997 TCGAAGACTGTTCATCTTAACTTAACATTTGAAGAGCCACAGAACTTGTATTTTGTG 1056
DB 783 TCGACGAGAGCTGGAAGAAAGCTCAGCTTGAAGAAATGCTGGAATCTGTGTGTGCTC 842
QY 1057 TCAGGTTTGCATCCCAATGATGAGACCTGTGAGAAAGTGATATTCAGTTAGGT 1116
DB 843 CAAGATTTTGTGCTGCGCAGATGTAATACGCTGTGATCAATGTGCGCACTTGCCT 902
QY 1117 TAGAATTCATCAACCTTAAGCTCCAGGCAACCTTATTACAG 1161
DB 903 AGCGTGTCTTATA---GTAATGATCCCGGCACTTTGATTTGCCG 944

Search completed: March 23, 2004, 11:07:11
Job time: 161 secs

1	1983	99.9	1985	9	US-09-890-813-5	Sequence 5, Appl1
2	1863.2	93.9	1953	9	US-09-890-813-7	Sequence 7, Appl1
3	1416.6	71.4	1533	12	US-10-425-114-5934	Sequence 5934, I
4	831.4	41.9	1115	12	US-10-425-114-3355	Sequence 3355, I
5	745.4	37.6	1780	12	US-10-425-114-6176	Sequence 6176, I
6	740.6	37.3	2882	12	US-10-424-559-28065	Sequence 28065, I
7	707.6	35.6	1710	9	US-09-838-842A-1744	Sequence 1744, A
8	707.6	35.6	1710	11	US-09-938-842A-1744	Sequence 1744, A
9	674.6	34.0	1658	9	US-09-890-813-15	Sequence 15, Appl1
10	655.4	33.0	1582	12	US-10-425-114-3869	Sequence 3869, I
11	609.2	30.6	1361	12	US-10-425-114-8335	Sequence 8335, I
12	555	28.0	565	9	US-09-890-813-1	Sequence 1, Appl1
13	533.2	26.9	2490	12	US-10-424-559-2060	Sequence 2060, I
14	515	25.9	1671	12	US-10-425-114-7821	Sequence 7821, I
15	470.4	23.7	1260	12	US-10-425-114-7451	Sequence 7451, I

[illegible]

Db 121 CTTCCGATACCCGAGCCGAGCTCGCGGTGCAAGAGGTTGTCAAATGGTGGTGCAGCTC 180
 Qy 181 CACACACCGCGGCGCAAGCGAGCGGAGCGGCGCGCTTGGGGCGCGCTTCT 240
 Db 181 CACACACCGCGGCGCAAGCGAGCGGAGCGGCGCGCTTGGGGCGCGCTTCT 240
 Qy 241 CGAGGCGCTCGGAGTGAAGGAGTTGGGGGATCAGCTCAGCGTGTATGAAGTTGGGG 300
 Db 241 CGAGGCGCTCGGAGTGAAGGAGTTGGGGGATCAGCTCAGCGTGTATGAAGTTGGGG 300
 Qy 301 GTCTCGGTGTGTGGCGCGCGAGGATGGCTGAGGTTGGCGCGCTCATCTGAGCTTCC 360
 Db 301 GTCTCGGTGTGTGGCGCGCGAGGATGGCTGAGGTTGGCGCGCTCATCTGAGCTTCC 360
 Qy 361 CGAGAGCGCGCGGTGTGTCTCTTCCATGGGGAACCAACCAACCTTCTCT 420
 Db 361 CGAGAGCGCGCGGTGTGTCTCTTCCATGGGGAACCAACCAACCTTCTCT 420
 Qy 421 TGGTGAAGAGAGCGAGTGGGAGTGAATGCAATGTTCTGAAATCGAAGTGGAA 480
 Db 421 TGGTGAAGAGAGCGAGTGGGAGTGAATGCAATGTTCTGAAATCGAAGTGGAA 480
 Qy 481 TATGTCAAAGCTCATATCAAGACGCTGATGAACTTGGACCTTCCAGNATCTGTA 540
 Db 481 TATGTCAAAGCTCATATCAAGACGCTGATGAACTTGGACCTTCCAGNATCTGTA 540
 Qy 541 TACAGCCTTATGAATGAGCACTATTGAAGTATCGTATGATGAAGAGCTGAC 600
 Db 541 TACAGCCTTATGAATGAGCACTATTGAAGTATCGTATGATGAAGAGCTGAC 600
 Qy 601 GCCTAGAGCAGTGAATCTGTTTCAATTTGAGAAATGCAATGTCACAGATTTTTC 660
 Db 601 GCCTAGAGCAGTGAATCTGTTTCAATTTGAGAAATGCAATGTCACAGATTTTTC 660
 Qy 661 TGGTATTGGAACAAATTCGTGTCAAGGACGAGCAATGATGATGATTTGTT 720
 Db 661 TGGTATTGGAACAAATTCGTGTCAAGGACGAGCAATGATGATGATTTGTT 720
 Qy 721 CATTAACACTGATGATTTGTTGTAAGCGATCTTGAAGCAACCTTCCGCTGTTGC 780
 Db 721 CATTAACACTGATGATTTGTTGTAAGCGATCTTGAAGCAACCTTCCGCTGTTGC 780
 Qy 781 GAAGAGCTTCAATGGGAGCTGATACAGATCCAGCATCTGTTGTTACTGGGTCT 840
 Db 781 GAAGAGCTTCAATGGGAGCTGATACAGATCCAGCATCTGTTGTTACTGGGTCT 840
 Qy 841 TGGAGAGGCTGGAATCTGCTGTGTAATCTTGAAGCGAGTGGTATGACTTGA 900
 Db 841 TGGAGAGGCTGGAATCTGCTGTGTAATCTTGAAGCGAGTGGTATGACTTGA 900
 Qy 901 TGGTACACCAATGGTGAAGCTTGGGACTGAGAAATTCAGGTATGAAAGATGTGA 960
 Db 901 TGGTACACCAATGGTGAAGCTTGGGACTGAGAAATTCAGGTATGAAAGATGTGA 960
 Qy 961 TGGTACTACTTGTGATCCAAATTCATCCACATGCAAGATGTTTCCATCTTAC 1020
 Db 961 TGGTACTACTTGTGATCCAAATTCATCCACATGCAAGATGTTTCCATCTTAC 1020
 Qy 1021 ATTGAAGAGGCGCAAGACTTGTATTGTTGGTGTCAAGTTTGCATCCATCAT 1080
 Db 1021 ATTGAAGAGGCGCAAGACTTGTATTGTTGGTGTCAAGTTTGCATCCATCAT 1080
 Qy 1081 GAGACCTGCTAGAGAGTGAATTCAGTGGGTGAAGATTCATCAACCTTAAGC 1140
 Db 1081 GAGACCTGCTAGAGAGTGAATTCAGTGGGTGAAGATTCATCAACCTTAAGC 1140
 Qy 1141 TCCAGGACCTTATTATCAAGACAAAGACATGTATANGTCTGTTGTTACTACTAG 1200
 Db 1141 TCCAGGACCTTATTATCAAGACAAAGACATGTATANGTCTGTTGTTACTACTAG 1200
 Qy 1201 CATAGGCTCAAGTCAATGTACTATGTTGACATTTGGAGCACTGGAGTGTGTC 1260

Db 1201 CATAGTCTCAAGTCAATGTACTATGTTGACATTTGTAGACATTTGTAGACTGCTTGTCA 1260
 Qy 1261 GTATGTTTTCGCAAGAGTATCAGGATTTGCTATATTAATGAATCTATATCTGT 1320
 Db 1261 GTATGTTTTCGCAAGAGTATCAGGATTTGCTATATTAATGAATCTATATCTGT 1320
 Qy 1321 GATGTTGTTGCTACCAAGTGAATTTAGTGTCTGCTCACTTATCCATCAAGATCTG 1380
 Db 1321 GATGTTGTTGCTACCAAGTGAATTTAGTGTCTGCTCACTTATCCATCAAGATCTG 1380
 Qy 1381 GATGAGGAGTGAATACAGCAGGAGTGAATTTGACATGTATGTTGAAGCTTGA 1440
 Db 1381 GATGAGGAGTGAATACAGCAGGAGTGAATTTGACATGTATGTTGAAGCTTGA 1440
 Qy 1441 ATATGATTTGTTGCTTACTTACAGCAGAGGCGTAAATTTCACTTATCGGAATGTGA 1500
 Db 1441 ATATGATTTGTTGCTTACTTACAGCAGAGGCGTAAATTTCACTTATCGGAATGTGA 1500
 Qy 1501 GCAATGCTCTGATATCAAGAAAGACGAGCGTGTCTGAGGAAAGTGGGGTTAATGT 1560
 Db 1501 GCAATGCTCTGATATCAAGAAAGACGAGCGTGTCTGAGGAAAGTGGGGTTAATGT 1560
 Qy 1561 TCAGATGATCTCGCAGAGAGCGTCAAAAGTTTACATGTCGCTGATGATGATGCGA 1620
 Db 1561 TCAGATGATCTCGCAGAGAGCGTCAAAAGTTTACATGTCGCTGATGATGATGCGA 1620
 Qy 1621 TGCAGAGCAGCTGAGAGCGCTTCAATCAGGCGTCTTGAAGACGATGTCTATCACA 1680
 Db 1621 TGCAGAGCAGCTGAGAGCGCTTCAATCAGGCGTCTTGAAGACGATGTCTATCACA 1680
 Qy 1681 AGTCAAGCGAGAACTTACTGTGGGCTGATCAAGTACGCTTGTGCTGCTCAAGCGT 1740
 Db 1681 AGTCAAGCGAGAACTTACTGTGGGCTGATCAAGTACGCTTGTGCTGCTCAAGCGT 1740
 Qy 1741 GTTATCTGTTATGATTCCTCAGCTGCGCTCAATGAAGCGCATTTGATCATGATC 1800
 Db 1741 GTTATCTGTTATGATTCCTCAGCTGCGCTCAATGAAGCGCATTTGATCATGATC 1800
 Qy 1801 ATGTTTCTTGAACAAAGTATGTTTCCAGGTTCTGACCAATGATCTGCAAACTGTGT 1860
 Db 1801 ATGTTTCTTGAACAAAGTATGTTTCCAGGTTCTGACCAATGATCTGCAAACTGTGT 1860
 Qy 1861 TTTCTGTTTGAAGCTGTTGAGACCAAGTACAGAGCTGCGACCGATTTGTCAAGAT 1920
 Db 1861 TTTCTGTTTGAAGCTGTTGAGACCAAGTACAGAGCTGCGACCGATTTGTCAAGAT 1920
 Qy 1921 GCGAAGCTGTGATTAATTCACACTGTCTTAATCAATATATTAATAAATTAACAA 1980
 Db 1921 GCGAAGCTGTGATTAATTCACACTGTCTTAATCAATATATTAATAAATTAACAA 1980
 Qy 1981 TATCT 1985
 Db 1981 TATCT 1985

RESULT 2
 US-09-890-813-7
 ; Sequence 7, Application US/09890813
 ; Publication No. US20020183486A1
 ; GENERAL INFORMATION:
 ; APPLICANT: E.I. du Pont de Nemours and Company
 ; TITLE OF INVENTION: Aspartate Kinase
 ; FILE REFERENCE: B01430 PCT
 ; CURRENT APPLICATION NUMBER: US/09/890, 813
 ; PRIORITY FILING DATE: 2001-08-02
 ; PRIOR APPLICATION NUMBER: 60/172944
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 7
 ; LENGTH: 1953
 ; TYPE: DNA
 ; ORGANISM: Zea mays

US-09-890-813-7

Query Match	93.9%	Score 1863.2	DB 9	Length 1953
Best Local Similarity	98.4%	Pred. No. 0		
Matches 1941	Conservative	0	Mismatches 8	Indels 23
				Gaps 5

QY	1	GACACAGAGACTCACAAGATGGCATCCACAGCGATCGGCTGCGCGCCCGCCGCTC	70
Db	1	GACACAGAGACTCACAAGATGGCATCCACAGCGATCGGCTGCGCGCCCGCCGCTC	60
QY	71	GTTCCGTCGATACCTCCGCGAGCTCTGACATGTTGAGGACTGCGCTGCTTCGATAC	130
Db	61	GTTCCGTCGATACCTCCGCGAGCTCTGACATGTTGAGGACTGCGCTGCTTCGATAC	120
QY	131	CGAACCGGAGCTCGCGGTGACAGAGGTTGTCAATGAGTGTGCGCCGACTCCACACACGT	190
Db	121	CGAACCGGAGCTCGCGGTGACAGAGGTTGTCAATGAGTGTGCGCCGACTCCACACACGT	180
QY	191	CGGCGCAAGACCGGACGCGCGGACCGCGCTCCTTGGGCGCCTGTTCCTGAGGCTC	250
Db	181	CGGCGCAAGACCGGACGCGCGGACCGCGCTCCTTGGGCGCCTGTTCCTGAGGCTC	240
QY	251	GGGATGAGAGGATTGGGGATCAGCTCAGCGGTGATGAAAGTTCCGAGGAGTCTCGAGT	310
Db	241	GGGATGAGAGGATTGGGGATCAGCTCAGCGGTGATGAAAGTTCCGAGGAGTCTCGAGT	300
QY	311	TGTCGCGCCGAGAGATGGCTGAGGTGCGCGGCTCATCTGACGTTCCCGAGAGCC	370
Db	301	TGTCGCGCCGAGAGATGGCTGAGGTGCGCGGCTCATCTGACGTTCCCGAGAGCC	360
QY	371	CCCGTGTGTTCTCTCTGACATGGGAGAAAACACCAACACTTCCTCTGTGAGAG	430
Db	361	CCCGTGTGTTCTCTCTGACATGGGAGAAAACACCAACACTTCCTCTGTGAGAG	420
QY	431	AAGCAGTAGGAGTGTGAGTTATCCATGTTCTGAATCGAAGAGTGAATATGTCAA	490
Db	421	AAGCAGTAGGAGTGTGAGTTATCCATGTTCTGAATCGAAGAGTGAATATGTCAA	480
QY	491	AGGCTCATATCAAGCGGTGATGATCTGAGCTCCAAAGATCTGATACAGCCCT	550
Db	481	AGGCTCATATCAAGCGGTGATGATCTGAGCTCCAAAGATCTGATACAGACAT	539
QY	551	T-----ATGAATCGAGCAACTATTTGAAGGATCGCTATGATGAAGAGCTGACGCTAG	606
Db	540	GCTAGATGAATCGAGCAACTATTTGAAGGATCGCTATGATGAAGAGCTGACGCTAG	599
QY	607	GACACATGACTACCTGTTTCAATTTGGAGATGCAATGCCACCGAATTTTCTGCTTA	666
Db	600	GACACATGACTACCTGTTTCAATTTGGAGATGCAATGCCACCGAATTTTCTGCTTA	659
QY	667	TTTGAACAAATTCGTGTCAAGGCACGGACAGTATGCGCATTTGATATTGTTCTTAC	726
Db	660	TTTGAACAAATTCGTGTCAAGGCACGGACAGTATGCGCATTTGATATTGTTCTTAC	719
QY	727	AATGATGAAATTTGGTAATGCGGATCTTTGAAGCAACTATCCGCTGTGCGAGAG	786
Db	720	AATGATGAAATTTGGTAATGCGGATCTTTGAAGCAACTATCCGCTGTGCGAGAG	779
QY	787	ACTTCATGAGGACTGGATACAGATCCAGCATCTGTTGTAATCTGGTTCCTTGAGAA	846
Db	780	ACTTCATGAGGACTGGATACAGATCCAGCATCTGTTGTAATCTGGTTCCTTGAGAA	839
QY	847	GGGCTGAAATCTGCTGCTGTAACTTAATTTAGCGCGAGGTGTGTGTACTTGAATCTAC	906
Db	840	GGGCTGAAATCTGCTGCTGTAACTTAATTTAGCGCGAGGTGTGTGTACTTGAATCTAC	899
QY	907	AACCATTTGTAAGCCTTGGGACTGAGAAATTCAGGTATGAGAAAGATGTATGATGT	966
Db	900	AACCATTTGTAAGCCTTGGGACTGAGAAATTCAGGTATGAGAAAGATGTATGATGT	959
QY	967	ACTTACTTGATGATCAAAATATCTACCCACAGCAAGATCTGTCATTAATTAATTTGA	1022
Db	960	ACTTACTTGATGATCAAAATATCTACCCACAGCAAGATCTGTCATTAATTAATTTGA	1011

Qy	1027	AGAGGCGACAGAACTGGTTATTTTGGTGTCAAGCTTTTGATCCCAATCGATCAGAC	1086
Db	1020	AGAGGCGACAGAACTGGTTATTTTGGTGTCAAGCTTTTGATCCCAATCGATCAGAC	1079
Qy	1087	TGCTAGAGAAAGTGATATTCAGTTAGGGTTAAGAAATCATACAACTTAAAGCTCCAG	1146
Db	1080	TGCTAGAGAAAGTGATATTCAGTTAGGGTTAAGAAATCATACAACTTAAAGCTCCAG	1139
Qy	1147	CACCCCTATTACCAAGCAAGACATGATTAANGTGTGGTTGTACTACTAGCATAGT	1206
Db	1140	CACCCCTATTACCAAGCAAGACATGATTAANGTGTGGTTGTACTACTAGCATAGT	1193
Qy	1207	GCTCAAGTCMAATGTCTACTATGTTGACATGTGTAGCACTCGGATGTCTTGTCAATATG	1266
Db	1194	GCTCAAGTCMAATGTCTACTATGTTGACATGTGTAGCACTCGGATGTCTTGTCAATATG	1253
Qy	1267	TTTTTCGGAGAGGGATCAGGATTTTGCTATATTGGAAGATCTATGATATCTGTGGATTG	1326
Db	1254	TTTTTCGGAGAGGGATTTGCTATAT-----TTGAAGATCTATGATATCTGTGGATTG	1307
Qy	1327	TGTTGTACCAAGTAGTAGTGTGTTGTGTGTCACTTGATCCATCAAAATCTGGATAG	1386
Db	1308	TGTTGTACCAAGTAGTAGTGTGTTGTGTGTCACTTGATCCATCAAAATCTGGATAG	1367
Qy	1387	GGAACTGATACACGACGGCAATGAACTTGACACTGATAGTAGTAAGAGCTTGAAATATGC	1446
Db	1368	GGAACTGATACACGAG-----GAACCTGACACTGATAGTAGTAAGAGCTTGAAATATGC	1421
Qy	1447	AATGTTGCTGTACTCTTCAGCAGAGGGCGGATTAATTTCACTTATCGAAATGTGGACAAATC	1506
Db	1422	AATGTTGCTGTACTCTTCAGCAGAGGGCGGATTAATTTCACTTATCGAAATGTGGACAAATC	1481
Qy	1507	GTTCTCGATACTAAGAAAAGCGGACGTGTGTGTGAAGAAAAGTGGGTTAATGTTCAAT	1566
Db	1482	GTTCTCGATACTAAGAAAAGCGGACGTGTGTGTGAAGAAAAGTGGGTTAATGTTCAAT	1541
Qy	1567	GATCTGCGAAGAGCGTCMAAGGTTAACATGTGCGTGTATGTCATGATAGCGATGCAAA	1626
Db	1542	GATCTGCGAAGAGCGTCMAAGGTTAACATGTGCGTGTATGTCATGATAGCGATGCAAA	1601
Qy	1627	GGCACTCGTAAGAGCCCTTTCATCAGGGGCTCTTTGAAGACGATGCTATCCACAGTGA	1688
Db	1602	GGCACTCGTAAGAGCCCTTTCATCAGGGGCTCTTTGAAGACGATGCTATCCACAGTGA	1661
Qy	1687	AGCGAGAACCTACCTCGTGGGCTGATCAAGATAGGCTTTGCTGGGTCCAGCGGTATATC	1746
Db	1662	AGCGAGAACCTACCTCGTGGGCTGATCAAGATAGGCTTTGCTGGGTCCAGCGGTATATC	1721
Qy	1747	TGTTATATAGTTCCCACTGCTCCATGAAACGGCATGGGCAATTGGACATTTGATCATGTTT	1806
Db	1722	TGTTATATAGTTCCCACTGCTCCATGAAACGGCATGGGCAATTGGACATTTGATCATGTTT	1789
Qy	1807	TGCTTAACAACAGTATGTCTTCCAGGTTTCCAGCAATGACTGCAAACTGTGTTCTGT	1866
Db	1782	TGCTTAACAACAGTATGTCTTCCAGGTTTCCAGCAATGACTGCAAACTGTGTTCTGT	1841
Qy	1867	TTTTAAGACTGTTTGCAGACACCAAGTAGACTGCGAGCACGATTTGTCAACAAATGGCAAG	1922
Db	1842	TTTTAAGACTGTTTGCAGACACCAAGTAGACTGCGAGCACGATTTGTCAACAAATGGCAAG	1901
Qy	1927	CCTGTATATTAATTCCAACTGTCTATCAATATATATATTAACATTATC 1978	
Db	1902	CCTGTATATTAATTCCAACTGTCTATCAATATATATATTAACATTATC 1953	

```
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 5934
/ LENGTH: 1533
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: 700553382_FLI
US-10-425-114-5934

Query Match      71.4%; Score 1416.6; DB 12; Length 1533;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 1509; Conservative 0; Mismatches 9; Indels 27; Gaps 6;

QY 449 GTATTCATGTTTCTGAAATCGAAGATGGAATATGTCAAAAGCTTACATATCAGACG 558
DB 1 GTATTCATGTTTCTGAAATCGAAGATGGAATATGTCAAAAGCTTACATATCAGACG 60
QY 509 GTGATGAACTGGACCTTCCAGNATCTGTATTAAG----CCTTTATGAACTGAGCA 564
DB 61 GTGATGAACTGGACCTTCCAGNATCTGTATTAAG----CCTTTATGAACTGAGCA 119
QY 565 ACTATTGAAAGGTATCGCTATGATGAAAAGAGCTGACGCTTGAAGCAAGTGAACCTTGT 624
DB 120 ACTATTGAAAGGTATCGCTATGATGAAAAGAGCTGACGCTTGAAGCAAGTGAACCTTGT 179
QY 625 TTCAATTTGGAAGATGATGTCACCAAGATTTTCTGCTATTGTAACAAATTCGTGT 684
DB 180 TTCAATTTGGAAGATGATGTCACCAAGATTTTCTGCTATTGTAACAAATTCGTGT 239
QY 685 CAAGGACGCGAGTATGACGATTTGATTTGTTTCACTAACAATGATGATTTGTTAA 744
DB 240 CAAGGACGCGAGTATGACGATTTGATTTGTTTCACTAACAATGATGATTTGTTAA 239
QY 745 TCGGATATCTTGAAGCAACTATTCCTGCTTTGCGAAGAGATCTTATGGGAGCTGGAT 804
DB 300 TCGGATATCTTGAAGCAACTATTCCTGCTTTGCGAAGAGATCTTATGGGAGCTGGAT 359
QY 805 ACAGGATCCAGGATACCTGTTGTTACTGAGGTTCTTGGGAAAGGCTGGAATCTGCTGC 864
DB 360 ACAGGATCCAGGATACCTGTTGTTACTGAGGTTCTTGGGAAAGGCTGGAATCTGCTGC 419
QY 865 TGTACTACTTTAGGCGCAGGTGTGATGATCTTGACTGCTCAACCATTTGTTAAAGCCTT 924
DB 420 TGTACTACTTTAGGCGCAGGTGTGATGATCTTGACTGCTCAACCATTTGTTAAAGCCTT 479
QY 925 GGGACTGAGAGAAATCAGGTATGGAAGAAGTGTGATGATGATGATGATGATGATGAT 984
DB 480 GGGACTGAGAGAAATCAGGTATGGAAGAAGTGTGATGATGATGATGATGATGATGAT 539
QY 985 TATCTACCAATGCAAGAGCTGTTCCATCTTAACTTTAAAGAGGCGACAGAACTTGC 1044
DB 540 TATCTACCAATGCAAGAGCTGTTCCATCTTAACTTTAAAGAGGCGACAGAACTTGC 599
QY 1045 TTATTTTGGTGTGAGTTTGTGATCCCAATCGATGAGAGCTGCTAGAGAGGTTAT 1104
DB 600 TTATTTTGGTGTGAGTTTGTGATCCCAATCGATGAGAGCTGCTAGAGAGGTTAT 659
QY 1105 TCCAGTTAGGTTAAGATTCATCAACCTTAAAGCTCCAGGACCTTATATCCAGACA 1164
DB 660 TCCAGTTAGGTTAAGATTCATCAACCTTAAAGCTCCAGGACCTTATATCCAGACA 719
QY 1165 AAGAGCATGATGAAGGTCGTGTTGATTAATGATGATGATGATGATGATGATGATGAT 1224
DB 720 AAGAGCATGATGAAGGTCGTGTTGATTAATGATGATGATGATGATGATGATGATGAT 773

/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 33565
/ LENGTH: 1115

RESULT 4
US-10-425-114-33565
/ Sequence 33565, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 33565
/ LENGTH: 1115

QY 1225 TATGTTGACATGTTGAGCACTGCGATGCTTGTGATGATGATGATGATGATGATGAT 1284
DB 774 TATGTTGACATGTTGAGCACTGCGATGCTTGTGATGATGATGATGATGATGATGATGAT 833
QY 1285 AGGATTTGCTATTTGAAGATCTAGTATATCTGCGATGATGATGATGATGATGATGATGAT 1344
DB 834 TGTATAT-----TTGAAGATCTAGTATATCTGCGATGATGATGATGATGATGATGATGAT 887
QY 1345 TAGTGTTCCTGTGTCACTTGTATCCATCAAAAGATCTGAGATGAGAACTGATACAGAGGC 1404
DB 888 TAGTGTTCCTGTGTCACTTGTATCCATCAAAAGATCTGAGATGAGAACTGATACAGAGGC 945
QY 1405 AAGTGAATTTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1464
DB 946 ----GAACTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1001
QY 1465 GCAGAGGCGCATTAATTTCACTTATCGGAAATGAGCAATCGCTCTGTATCTAGTAAATA 1524
DB 1002 GCAGAGGCGCATTAATTTCACTTATCGGAAATGAGCAATCGCTCTGTATCTAGTAAATA 1061
QY 1525 GACGGAGCTGTGCTGAGGAAAGTGGGTTAATGTTCAATGATCTTCCAGAGAGCTGC 1584
DB 1062 GACGGAGCTGTGCTGAGGAAAGTGGGTTAATGTTCAATGATCTTCCAGAGAGAGCTGC 1121
QY 1585 AAAGGTTAATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1644
DB 1122 AAAGGTTAATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1181
QY 1645 TCATGACGCTTCTTTGAAGCGATGCTTATCAAAATGGAACCGGAGAACTTATGCT 1704
DB 1182 TCATGACGCTTCTTTGAAGCGATGCTTATCAAAATGGAACCGGAGAACTTATGCT 1241
QY 1705 GGGCTGATCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1764
DB 1242 GGGCTGATCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1301
QY 1765 GCTTCATGAAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1824
DB 1302 GCTTCATGAAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1361
QY 1825 CTTCAGGTTCTGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1884
DB 1362 CTTCAGGTTCTGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1421
QY 1885 CACCA-----GTGAGCTGAGACACCGATGATGATGATGATGATGATGATGATGAT 1940
DB 1422 CACCAAGCTGAGCTGAGACACCGATGATGATGATGATGATGATGATGATGATGATGAT 1481
QY 1941 CCAAGTGTCTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1985
DB 1482 CCAAGTGTCTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1526
```

TYPE: DNA
 ORGANISM: Zea mays
 FEATURE:
 OTHER INFORMATION: Clone ID: UC-ZMFLMO17134C03_FLI
 US-10-425-114-33565

Query Match 41.9%; Score 831.4; DB 12; Length 1115;
 Best Local Similarity 97.0%; Pred. No. 3.9e-250;
 Matches 897; Conservative 0; Mismatches 6; Indels 22; Gaps 4;

1058 CAGGTTTGCATCCACATCGATGAGACCTGCTAGAGAGAGTATATTCAGTGGGTT 1117
 209 CAGGTTTGCATCCACATCGATGAGACCTGCTAGAGAGAGTATATTCAGTGGGTT 268
 1118 AAGATTTCATACACCTTAAGCTCCAGGACCCCTTATTACAGACAAAGAGACATGAT 1177
 269 AAGATTTCATACACCTTAAGCTCCAGGACCCCTTATTACAGACAAAGAGACATGAT 328
 1178 AAGGCTCTGTTTACTACTAGCATAGTGTCTCAAGTCAATGTCTATGTGGACAT 1237
 329 AA-----GGTTTACTACTAGCATAGTGTCTCAAGTCAATGTCTATGTGGACAT 382
 1238 GTGAGCACTCGGATGCTTGGTCAATAGTGTCTTCTGCAAGGATTCAGGATTTGCTAT 1297
 383 GTGAGCACTCGGATGCTTGGTCAATAGTGTCTTCTGCAAGGATTTGCTATAT----- 437
 1298 ATTGAAGATCTATGATCTGATCTGATGATGTGTGTGTCTACAGAGAGATGATGTGTGTG 1357
 438 -TTGAAGATCTATGATCTGATCTGATGATGTGTGTGTCTACAGAGAGATGATGTGTGTG 496
 1358 TCACTTGATTCATCAAGATCTGAGTACGAGTAACTGATACAGAG-----GAACCTTAC 550
 497 TCACTTGATTCATCAAGATCTGAGTACGAGTAACTGATACAGAG-----GAACCTTAC 550
 1418 CAGTGTGTAAGAGCTTGAGAAATAGCAATGTGTCTCTACTCAGAGAGGCGCATTA 1477
 551 CAGTGTGTAAGAGCTTGAGAAATAGCAATGTGTCTCTACTCAGAGAGGCGCGCATTA 610
 1478 ATTCACTTATCGGAATGTGAGCAATGTCTCTGATATCTAGAAAGAGCGGACGTGTG 1537
 611 ATTCACTTATCGGAATGTGAGCAATGTCTCTGATATCTAGAAAGAGCGGACGTGTG 670
 1538 CTGAGGAAAGTGGGTTATGTTTCAGATGATCTGCAAGAGCGGCAAGGTTACATG 1597
 671 CTGAGGAAAGTGGGTTATGTTTCAGATGATCTGCAAGAGCGGCAAGGTTACATG 730
 1598 TCGCTGATGTCATGATAGCGATGCAAGAGCACTGTAGAAAGCCCTTCATCAGGCGTT 1657
 731 TCGCTGATGTCATGATAGCGATGCAAGAGCACTGTAGAAAGCCCTTCATCAGGCGTT 790
 1658 TTGGAAGACATGTCTTATACAGAGTCAAGAGCGGAAACCTACTGCTGGGCTGATCAAG 1717
 791 TTGGAAGACATGTCTTATACAGAGTCAAGAGCGGAAACCTACTGCTGGGCTGATCAAG 850
 1718 TAGGCTTTGCTGGGTCAGGCGGTATATCTGTTATAGATTTCCACTGCGCTCCATGAAG 1777
 851 TAGGCTTTGCTGGGTCAGGCGGTATATCTGTTATAGATTTCCACTGCGCTCCATGAAG 910
 1778 GCATGGGATTTGATCATTTATCATGTTTGTGTTGAACAAGATGTCTTCCAGGTTCTC 1837
 911 GCATGGGATTTGATCATTTATCATGTTTGTGTTGAACAAGATGTCTTCCAGGTTCTC 970
 1838 AGCAATGATGCAAACTGTGTCTGTTTGAAGTGTGTTGACAGACCA-----GTGA 1893
 971 AGCAATGATGCAAACTGTGTGTCTGTTTGAAGTGTGTTGACAGACCAAGATGTGA 1030
 1894 GCTGGAGACCGATTTGTCAACAAGATGCAAGCTGTGATATATTCGAATGTGTCTA 1953
 1031 GCTGGAGACCGATTTGTCAACAAGATGCAAGCTGTGATATATTCGAATGTGTCTA 1090
 1954 ATCAATATATATATATAACATTATC 1978
 1091 ATCAATATATATATAACATTATC 1115

RESULT 5
 US-10-425-114-6176
 ; Sequence 6176, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yungwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 6176
 ; LENGTH: 1780
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 70056108_FLI
 US-10-425-114-6176

Query Match 37.6%; Score 745.4; DB 12; Length 1780;
 Best Local Similarity 72.6%; Pred. No. 6.4e-223;
 Matches 1010; Conservative 0; Mismatches 367; Indels 15; Gaps 3;

284 GTGATGAATTCGGGGGCTCTCGGTGCTGCTGCGCCGAGAGATGCTGAGGTGCGCGG 343
 48 GTGATGAATTCGGGGGCTCTCGGTGCTGCTGCGCCGAGAGATGCTGAGGTGCTGAG 107
 344 CTGATCTGACGTTCCCGAGAGAGCGCCCGTGTCTCTCTCTGCTGCTGAGGAGAAACC 403
 108 CTGATCTGACGTTCCCGAGAGAGCGCCCGTGTGTCTCTCTCTGCTGCTGAGGAGAAACA 167
 404 ACCAACAACCTTCTCTCTGCTGAGAGAGAGGAGGAGTGTGAGTATTCAGTTTCT 463
 168 ACAACAACCTTCTCTCTGCTGAGAGAGAGGAGGAGTGTGAGTATTCAGTTTCTA 227
 464 GAAATGGAAGAGGAAATGCTCAAAAGCTTACATCAAGAGGTGATGAACCTTGA 523
 228 AGTATGAGAGAGCTTGTCTTATTAAGACCTGATCAAGAGCTGAGATCACTTGGT 287
 524 CTTCAGAGATCTGTAATCAAA--GCCCTTATGAACTGAGCAACTATGAAAGTATC 580
 288 GTGAGAGAGCTGTAATCAAAAGCATCTAGAGAGATGAGCAACTTCTGAAAGGAGTA 347
 581 GCTATGATGAAGAGAGCTGTAAGAGCATCAAGAGCTATTTAGTCTCTTGGAGATGC 640
 348 GCTATGATGAAGAGAGCTGTAAGAGCATCAAGAGCTATTTAGTCTCTTGGAGATGC 407
 641 ATGTCCACAGAGATTTTCTGCTTATTTGAACAAATTCGTGCAAGGACGCGAGTAT 700
 408 ATGTCCACAGAGATTTTCTGCTGATATCTTATTAAGAGTGTGCAAGGCTGCGCAATAT 467
 701 GAGGATTTGATTTGCTTCACTTCACTGATGAAATTTGGAATGAGGATATCTTGA 760
 468 GATGATTTGATTTGCTTCACTTCACTGATGAAATTTGGAATGAGGATATCTTGA 527
 761 GCAACTTATCTGCTGTTGCGAAGAGCTTCAAGGAGCTGATACAGATCCAGGATA 820
 528 GCAACTTATCTGCTGTTGCGAAGAGCTTCAAGGAGCTGATACAGATCCAGGATA 587
 821 CCGTGTGTTACTGAGTCTTGGGAGAGGCTGGAATGCTGCTGTAACCTTTAGGC 880
 588 GCAATGTTACTGAGTCTTGGGAGAGGCTGGAATGCTGCTGTAACCTTTAGGC 647
 881 CGAGTGTGATGATCTGATCAACATTTGTAAGCTTGGAGCTGAGGAAAT 940


```

Db      648 AGAGGGGAGGATTTGACGGCTACAACAATTGGGAAGCACTAGGGTTGCTGAGATC 707
QY      941 CAGGTATGGAAGATGTTGATGGTACTTGTGATGCCAAATATCTACCCACATGCA 1000
Db      708 CAGGTATGGAAGATGTTGATGGTACTTGTGATGCCAAATATCTACCCACAAAGCA 767
QY      1001 AAGACTGTTCCATCTTAAACATTTAAGAGGCCACAGAACTTGTCTATTTTGGTCTCAG 1060
Db      768 GAACCTGTTCTTATTTGACATTTGATGAGGCTGACAGAACTTGTCTTATTTGGTCTCAG 827
QY      1061 GTTTGATCCCAATTCGATGAGACCTGCTAGAGAGGTGATATTTCCAGTTAGGTTAG 1120
Db      828 GTTTCATCCACAGCTTATGAGACCTGCTAGAGAAATGATATTTCTGTTAGGTTAA 887
QY      1121 AATTATACCACTTAAAGCTCCAGGCACTCTTATTTACCAACAAGAGCATGATTAAN 1180
Db      888 AATTCTTACACCTTAAAGCTCCAGGTACTCTCATCACCAAGGCAAGATATGAGCAA- 946
QY      1181 GGTCTGTTGTAATACTAGCACTAGTCTCAAGTCAATGTCATATGTTGGAATGTTG 1240
Db      947 -----GGCAGTATTAACAAGCAATGTTTGAACCTAATGTGACATGTTGATATGCA 1001
QY      1241 AGCACTCGATGCTTGTGTCAGTATGATTTCTGCAAGGTTATCAGTATTTGCTATATT 1300
Db      1002 AGCACTCGATGCTTGTGTCAGTATGATTTCTGCTAAGTGT-----TTTCAATCTTT 1055
QY      1301 GAAGATCTATGATATCTGTGATTTGTGTCTACCACTGAAAGTTAGTTTCTGTCTCA 1360
Db      1056 GAAGATGTTGCAATCAAGTATGATTTGTGATGCTAAGTGAAGTCAAGTCTTCTTACA 1115
QY      1361 CTGTATCCATCAAAAGTCTGAGTAGAGGCACTGATACAGACGCAAGTGAATTTAGCAT 1420
Db      1116 CTGTATCCATCAAAAGTATGAGACGAGAGCTAATTTCAAGGCAAGTGAATTTAGCAT 1175
QY      1421 GTATGTTGAGAGCTTGAAGAAATAGCAATTTGTTCTCTTACTTCAAGAGGCGCATATT 1480
Db      1176 GTTGTAGAAAGAACTCAGAAATTCGCTGTGTAATCTCTCAAGAAATGATTCATATC 1235
QY      1481 TCACCTATCGAAATGTGAGCAATGCTCTGATACTAGAAAGACGGGAGCTGTCTG 1540
Db      1236 TCTCTCATGGAATTTTCAAGATCATCATTAATTTGGAAGGCTTTCCGTTCTT 1295
QY      1541 AGGAAAAGTGGGTTATTTCAATGATCTCGCAAGAGGCTCAAGGTTAAATGTCG 1600
Db      1296 CGAAACCTTGGGCTCACTGTGCAATGATCTCCAGGGTGCATCTAAAGTAAACATCTCA 1355
QY      1601 CTGATATGCATGATAGCGATGCAAAAGGCACTCGTAAAGGCTTCAAGGCTTCTTT 1660
Db      1356 TTGGTTGTAATGACGTGAACGACGAGCTGTGTGAGGCTTCTCCTCACTCAGCTTCTTT 1415
QY      1661 GAAGACGATGTC 1672
Db      1416 GAGAGTGAAGCTC 1427

```

RESULT 6

```

US-10-424-599-28065
; Sequence 28065, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424, 599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 28065
; LENGTH: 2882
; TYPE: DNA

```

```

; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MR1847_125343C.1
US-10-424-599-28065

Query Match      37.3%; Score 740.6; DB 12; Length 2882;
Best Local Similarity 72.3%; Pred. No. 3e-221; Indels 15; Gaps 3;
Matches 1007; Conservative 0; Mismatches 370

QY      284 GTGATGAAGTTCGGGGGCTCTCGTGTGCTGCGCCGAGAGATGAGTGAAGTGGCCGC 343
Db      918 GTCATGAAGTTCGGGGGCTCTCGTGTGCTGCGCTTGTGATGATGAAGAGTGGCTACC 977
QY      344 CTCATCTGACGTTCCCGAGAGAGCCCGTGTGCTTCTCTCCATTTGGGAAATCC 403
Db      978 CTTATATTGAAGTTCCTCGAGAGAGGCTTATTTGTTCTCTCTGCTATGGGAAATCA 1037
QY      404 ACCAACAACCTTCTCTGCTGAGAGAGGCAAGTGGTGTGAGTTATCCATGTTTCT 463
Db      1038 ACAACCAAGCTTTGCTGCTGAGAGAGAGAGAGAGAGAGTGTGATGATGATATCA 1097
QY      464 GAATTCGAAGTGAATATGCTCAAAAGCTTACATATCAAGCGGTGATGAATCTTGA 523
Db      1098 AGTATTTGAGAGCTTTGCTTATTAAGACCTGATCTAAGACTGTGATCAGCTGTGT 1157
QY      524 CTTCAGATCTGTAATACAA--GCCTTATGAACTGAGAGCACTATTGAAGGTATC 580
Db      1158 GTGAGCGATCTGTTATTTCAAGCATCTAAGAAATGAGCACTTCTGAAGGGATA 1217
QY      581 GCTATGATGAAGAGCTGACGCTTAGACCAAGTACTGCTTGTTCATTTGAGAAATGC 640
Db      1218 GCTATGATGAAGAGATGACTTAAAGAGCTCAGACATATTATGCTCTTTGAGAAATGC 1277
QY      641 ATGTCACAGAGATTTTTCGCTTATTGAAACAAATTTGCTGCAAGGACGAGCTAT 700
Db      1278 ATGTCACAGAGATTTTTCGCTTATTGAAACAAATTTGCTGCAAGGACGAGCTAT 1337
QY      701 GACGATTTGATATTGGTTTCATTACACTGATGAATTTGTTATGCGATATCTTAGAA 760
Db      1338 GATGATTTGATATTGGTTTCATTACACTGATGAATTTGTTATGCGATATCTTAGAA 1397
QY      761 GCAACCTATCCTGCTTGGGAGAGAGCTCATAGGGGACGATACAGATCCAGGGATA 820
Db      1396 GCACTTATCCAGCTGTTGCAAGAGATTAATGATGATTTGCTCTGATCTCTCAAT 1457
QY      821 CCTGTTGTTACTGAGTTCTTGGAGAGGCTGAAATCTGCTGCTGATTAATCTTAAAGC 880
Db      1458 GCAATTTGTTACAGGCTTCTTGGAAAGCCCGGAAATCATGTGACAGTACACATGGGT 1517
QY      881 CAGGTGTGATGACTTGTGCTTACCAACATTTGTAAACCTTTGGGACTGAGAGAAAT 940
Db      1518 AGAGGGGAGATGATTTGACAGCTAACCAATTTGGGAAACACTAGAGTTGCTGAGATC 1577
QY      941 CAGGTATGGAAGATGTTGATGCTGTACTTACTTGTGATCAATATCTACCCACATGCA 1000
Db      1578 CAGGTATGGAAGATGTTGATGCTGTACTTACTTGTGATCAATATCTACCCACAAAGG 1637
QY      1001 AAGACTGTTCCATCTTAAACATTTGAAGAGCCACAGAACTTGTCTTATTTTGGTCTCAG 1060
Db      1638 GAACCTGTTCTTATTTGACATTTGATGAGGCTGCAAACTACGTAATCTTGTGTCTCAG 1697
QY      1061 GTTTGATCCACATGATGAGACCTGTAGAGAGGTATTTCCAGTTAGGTTAG 1120
Db      1698 GTTTCATCCACAGCTTATGAGACCTGCAAGAGAGGTATTTCCAGTTAGGTTAA 1757
QY      1121 AATTATACAACTTAAAGCTCCAGGCACTTATTTACGACAAAGAGACATGATTAAN 1180
Db      1758 AATTCTTACAACTTAAAGCTCCAGGTACTCTCATCCCAAGGCAAGATATGAGCAAA- 1816
QY      1181 GGTCTGTTGTAATACTAGCACTAGTGTCTCAAGTCAATGCTATGTTGACATTTG 1240
Db      1817 -----GGCAGTATTAACAAGCATTTTGAACGTAATGTGACATGTTGATATAGCA 1871

```

QY 1241 AGCACTCGATGCTGGTGCAGTATGGTTTCTGCGAAGGGTATCAAGTATTTGCTATAT 1300
 DB 1872 AGCACTCGATGCTGGTGCAGTATGGTTTCTGCGAAGGGTATCAAGTATTTGCTATAT 1925
 QY 1301 GAAGATCTATGATATCTGTGATTTGTGTCTACAGTGAAGTATGTTCTGTGCA 1360
 DB 1926 GAAGATCTATGATATCTGTGATTTGTGTCTACAGTGAAGTATGTTCTGTGCA 1985
 QY 1361 CTGATCCATCAAGATCTGAGTGAAGGAACTGATACAGCAGGCAAGTGAATTCACCAT 1420
 DB 1986 CTGATCCATCAAGATCTGAGTGAAGGAACTGATACAGCAGGCAAGTGAATTCACCAT 2045
 QY 1421 GTAGTGAAGAGTGGAGAAATAGCAATGTTGTGTCTACTGACAGCAGGCAAGTAT 1480
 DB 2046 GTTGTGAAGAACTGAGAAATCGCTGTGTGAATCTCTACAGAAATGATTCATCATC 2105
 QY 1481 TCATTAATCGAAATGTGAGCAATCGTCTGATATCAAGAAAGCAGGACGTGTGCTG 1540
 DB 2106 TCTCTATGAAATGTTCTGAGATCATCATATATTGGAAGGCTTTCGGTGTCTT 2165
 QY 1541 AGGAAAGTGGGGTTATGTTCAATGATCTTGCAAGAGCCTCAAGATTAAGTGTG 1600
 DB 2166 CGAACCTTGGGCTCCTGTGCAATGATCTCCAGGGTGCATCTAAGTGAACATCTCA 2225
 QY 1601 CTGATATCCATGATAGCATGCAAGGCACTCGTAGAAGCCCTTCATCAGGCGTCTT 1660
 DB 2226 TTGGTTGTAATGACAGTGAACGACAGCAGTGTGTGAGGGCTCTCCTCAGCGCTTCTT 2285
 QY 1661 GAAGACGATGTC 1672
 DB 2286 GAGAGTGAGCTC 2297

RESULT 7
 US-09-938-842A-1744
 ; Sequence 1744, Application US/09938842A
 ; Patent No. US20020160378B1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Keps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRIPI300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; PRIOR FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 1744
 ; LENGTH: 1710
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; US-09-938-842A-1744

Query Match 35.6%; Score 707.6; DB 9; Length 1710;
 Best Local Similarity 70.5%; Pred. No. 4,9e-211;
 Matches 1008; Conservative 0; Mismatches 401; Indels 21; Gaps 4;

QY 254 ATGAGGAGATTTGGGGGATCAAGCTCAAGCTGTGATGAAGTGGGGGGTCTCTGGTGTG 313
 DB 235 ACCGAGGTAGATAGAGGGTATCAAGTGTGATGAAGTGGGATCTTCGGTGTG 294
 QY 314 TCGGCGCGAGATGAGTGTGAGGCTCACTGATCGTTCGCCGAGAGGCGCC 373
 DB 295 TCGGCGCGAGATGAGTGTGAGGCTCACTGATCGTTCGCCGAGAGGCGCC 354
 QY 374 GTGCGTGTCTCTGTGCGATGGGAAACCAACCAACACTTCTCTGTGCGAGAGAG 433

DB 355 GTCAATGTTCTCTCTGCTATGAGGAAATCAACCAACTCTTGTCTGGGAGAGAG 414
 QY 434 GCAGTAGGGTGTGAGTATTCATGTTCTGAAATCGAAGATGGAATATGTGCAAGC 493
 DB 415 GCGTAGTGTGGTGTCTTATGATCATGTAGATGAGAGTGAACATTAATAAGAA 474
 QY 494 CTACATATCAAGCGGTGATGACTGAGCT---TCCAGGATCTGATTAACAAGCTT 550
 DB 475 TTGATATGAGACGGTGAAGAGCTCAACTGATATCTCTGTATTTGACCTATTG 534
 QY 551 TATGAATCGAGCACTATTGAAGATTCGCTATGATGAAGAGAGCTGAGCCTAGAC 610
 DB 535 GAGGAATCGAGCAACTCTGAAAGGCTTCCATGATGAAGGATGACATCTCGAAC 594
 QY 611 AGTGAATCTGTTTCTTTGAGAAATGATGATTCACAGAGATTTTCTGCTTATTG 670
 DB 595 AGAGATTACTTATGCTCTTTGAGAGTGTGTCTACAGGATTTTGTGCTTATCT 654
 QY 671 AACAAATTCGTGTAAGGACGAGCAAGATGACGATTTGATATGTTGTTTCAATTAAC 730
 DB 655 AATACATCGGTGTGACGAGCAGCAATGATGATGATTTGAATGTTTCAATTAAC 714
 QY 731 GATGAATTTGATGCGGATATCTTGAAGCAACTATCTGCTGTGCAAGAGACT 790
 DB 715 GATGATTTCAAAATGCGGATATCTGAGGCACTATCAGCTGTGCAAGAGATTA 774
 QY 791 CATGGGAGCTGGATACAGATTCACGATACCTGTTGTAATGAGGCTCTTGGGAGAG 850
 DB 775 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 834
 QY 851 TGAATATGTTGCTGTATCTTAAAGCCAGGCTGATGATGATGATGATGATGATG 910
 DB 835 TGAATATGTTGCTGTATCTTAAAGCCAGGCTGATGATGATGATGATGATGATG 894
 QY 911 ATTGTTAAAGCTTGGGCTGAGAGAAATCAGTATGGAAGATGTTATGTTGTT 970
 DB 895 ATTGTTAAAGCTTGGGCTGAGAGAAATCAGTATGGAAGATGTTATGTTGTT 954
 QY 971 ACTTGTATCCAAATATCTTACCCATGCAAGGCTGTTCAATCTTAACATTTGAAG 1030
 DB 955 AATGTAAGCTTATCTTATTAAGGAGCTACACAGATCAATCTGACATTTGAGAA 1014
 QY 1031 GCCAAGAACTGCTTATTTGGTGTGAGTTTGTGATGATGATGATGATGATGATG 1090
 DB 1015 GAGCGGAGCTAGCTTATTTGGTGTGAGTTTGTGATGATGATGATGATGATGATG 1074
 QY 1091 AGAGAAAGTATATCCAGTTAGGTTAAGATTCATCAACCTTAAGCTCCAGGAC 1150
 DB 1075 AGAGAAAGTATATCCAGTTAGGTTAAGATTCATCAACCTTAAGCTCCAGGAC 1134
 QY 1151 CTATATTAACAGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1210
 DB 1135 ATCATCATCAAAACAGAGATGACCAAGG-----TATTCAAAGACATTTGCTG 1188
 QY 1211 AAGTCAAAATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1270
 DB 1189 AAGTCAAAATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1248
 QY 1271 CTGGCAAGGATATCAAGTATTTGCTATTTGAAGATCTATGATATCTGTGATTTG 1330
 DB 1249 CTGGCAAGGATATCAAGTATTTGCTATTTGAAGATCTATGATATCTGTGATTTG 1302
 QY 1331 GCTACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1390
 DB 1303 GCTACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1262
 QY 1391 CTGATACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1450
 DB 1363 CTGATACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1416
 QY 1451 GTTGTGTATCTTCAAGAGAGGCGATATTTCACTTATCGGAAATGAGCAATCTCT 1510

Publication No. US20020183486A1
GENERAL INFORMATION:
APPLICANT: E.I. du Pont de Nemours and Company
TITLE OF INVENTION: Aspartate Kinase
FILE REFERENCE: B61430 PCT
CURRENT APPLICATION NUMBER: US/09/890,813
CURRENT FILING DATE: 2001-08-02
PRIORITY APPLICATION NUMBER: 60/172944
PRIORITY FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Microsoft Office 97
SEQ ID NO 15
LENGTH: 1658
TYPE: DNA
ORGANISM: Trilicium aestivum
US-09-890-813-15

Query Match 34.0%; Score 674.6; DB 9; Length 1658;
Best Local Similarity 72.8%; Pred. No. 1.2e-200;
Matches 930; Conservative 0; Mismatches 326; Indels 21; Gaps 4;

QY 425 GGAGAGAGGAGCTAGGCTGTGATTCATTCATTTCTGGAATGGAAGAGTGAATG 484
DB 4 GGACAGAGGGGCTGAGCTGGCGCCCAAGGCGTGGAAATCTACAGAGCTGGCGCTC 63
QY 485 GTCAAAAGCCTACATATCAAGAGCGGTGATGAATCTGCAAGNATCTGTATACA 544
DB 64 ATCAAGAGAGCTCATCTCAGAGCAATCATGATGAGCTTGGCTGATGAGCTCATGTTTCA 123
QY 545 AGCCTTT--ATGAATCGAGCACTATTTGAAGATGCTATGATGAAGAGTGAAG 601
DB 124 GGTATTTTGGACAGATGGAGCAATGCTCAAGGGGTGTGATATGAAGAGTGAAT 183
QY 602 CCTAGAGCAGTACTACCTGTTTCTTATTTGGAATGATGATGATGATGATGATGAT 661
DB 184 CTTAGAGACAGATTAATCTGTTTCTTGTGATGATGATGATGATGATGATGATGAT 243
QY 662 GCTTATTTGAACAAATTCGTTCANAGCAAGGAGATGAGCATTTGATTTGTTTC 721
DB 244 GCATATTTGAATTAATCAGAGGAAGAGGACAGCATGATGATGATGATGATGATGAT 303
QY 722 ATTAACAATGATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 781
DB 304 ATTAACAATGATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 363
QY 782 AAGAGATTCATGAGGAGCTGATATCAAGATTCAGATGATGATGATGATGATGATGAT 841
DB 364 AAGAGCTCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 423
QY 842 GGAAGAGGCTGGAATTCGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 901
DB 424 GGAAGAGGATGGAATTCGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 483
QY 902 GCTACAACATTTGATTAAGCTTGGATGAGAGAAATTCAGATGATGATGATGATGAT 961
DB 484 GCTACAACATTTGATTAAGCTTGGATGAGAGAAATTCAGATGATGATGATGATGATGAT 543
QY 962 GGTGTACTTATCTGTATTCATAATCTACCAATGAGCAAGAAATCTGTTCCATTAACA 1021
DB 544 GGTGTGTGAGCGTGTATTCATAATCTACCAATGAGCAAGAAATCTGTTCCATTAACA 603
QY 1022 TTGAAGAGGCAAGAACTTGTCTATTTTGTCTCAGATTTTGTGATCCATCAATGATG 1081
DB 604 TTGTATGAGCAGAGTGAATTTGCTTATTTTGTGTCAGAGTTTGTGATCCCATATCAG 663
QY 1082 AGACCTCTAGAGAGGTATTTTCAAGTTAAGGTTAAGATTCATCAACCTTAAGCT 1141
DB 664 CGACCAAGCAGGAGAGGTGTATCCAGATTCAGATGAAGAACTCATTAACGTATGCA 723
QY 1142 CGAGGCAACCTTATTAACAGAAAGAGATGATGATGATGATGATGATGATGATGATGAT 1201
DB 724 CTGTGACTGTGATCACTAATAACAGAGATTCGCAAGAG-----CATATTAACGAC 777

QY 1202 ATAGTGTCAAGTCAAAATGTCATATGTTGAGCATTTGAGCACTCGGATGTTGAG 1261
DB 778 ATTTGCTGAAATCAAAATATTTCCATGCTGATATGAGCAACAAGATGCTGGACAG 837
QY 1262 TATGTTTTCGCAAGGGATTCAGATATTTGCTATATTTGAAGATCTATGATTCGTG 1321
DB 838 TATGCTTTCAGAAAGGTCTTCATATAT-----TTGAAGATTTGGGATCTCTGTT 891
QY 1322 GATTTGTTGCTACAGAGGAAGTATGTTTCTGTGCTACCTGATCCATCAAGATCTGG 1381
DB 892 GATTTGTTGCTACAGAGGAAGTATGTTTCTGTGCTACCTGATCCATCAAGATCTGG 951
QY 1382 AGTAGGAATGATACAGCAGGCAAGTGAATCTGACATGATGATGATGATGATGATGAT 1441
DB 952 AGTCGTGAATGATCCAGAG-----GAGCTTGTATGATGATGATGATGATGATGATGAT 1005
QY 1442 ATAGCAATTTGCTGCTACTTACAGAGAGGCGATTAATTTACTTATCGGAAATGAGAG 1501
DB 1006 ATTTGCGTTGTTCAATCTCTACAGCAGATCAATCATTTCTGATAGGAAATGTCAG 1065
QY 1502 CAATGCTCTGTATCTAGAAAAGACGAGAGTGTGCTGAGAAAAGTGGGTTATGTT 1561
DB 1066 AGATGCTCTGTATCTTGAAGAGGCGTTCAATGTTCTACAGCAAAATGCTTAACTTT 1125
QY 1562 CAGATGATTCGCAAGAGAGCGTTCAAGGTTAATGATGCTGATGATGATGATGATGAT 1621
DB 1126 CAGATGATTTCCCAAGAGGCGCTCCAAAGGTGAATCTCTGTTGTTGAAGTGAAGAG 1185
QY 1622 GCAAGAGCACTGTGAAGCCCTTCATCAGAGGCTTCTTGAAGAGATGCTTACAA 1681
DB 1186 GCAAGAGCACTGTGAAGCCCTTCATCAGAGGCTTCTTGAAGAGATGCTTACAA 1245
QY 1682 GTCAAGCGAGAACTT 1698
DB 1246 GTAGAGGAAGCGAGACT 1262

RESULT 10
US-10-425-114-3869
Sequence 3869, Application US/10425114
Publication No. US2004003488A1
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 3869
LENGTH: 1582
TYPE: DNA
ORGANISM: Zea mays
OTHER INFORMATION: Clone ID: 700343285_FLI
US-10-425-114-3869

Query Match 33.0%; Score 655.4; DB 12; Length 1582;
Best Local Similarity 72.3%; Pred. No. 1.3e-194;
Matches 912; Conservative 0; Mismatches 328; Indels 21; Gaps 4;

QY 456 ATGTTTCTGAATCGAAGGTGATATGTTCAAAAGCTTCATATCAAGCGGTGATG 515
DB 2 AGCGTTCGAAATCCCTGAGCTGCTGATCAAGAGATTCATCTTGAAGCGGTGATG 61
QY 516 AACTGAGCTTCCAGNATCTGTAATPACAGCTTT--ATGAAGTGAAGCACTATTTGA 572
DB 62 AGCTGGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121

QY 573 AAGGATTCGCTATGATGAAAGAGCTGACGCTAGAGCAAGTACTGCTGTTTCATTTG 632
 Db 122 AAGGAGTTGCTATGATGAAAGAGCTGACGCTAGAGCAAGATTTACTGTTTCTTTG 181
 QY 633 GAGAAATGATGCCAGAGATTTTTCGCTTATTTGAACAAAATTCGTCAAGGAC 692
 Db 182 GTGAATGATGTCAGACAGAAATTTTGTGCAATTTTGAACAAAATTCGTCAAGGAC 241
 QY 693 GGCAGTATGACGATTTGATTTGTTTCATTTACATGATGAATTTGGTAATGCGATA 752
 Db 242 GGCAGTATGATGATTTGATTTGTTTCATTTACACCGATGATTTTCAATGCGGATA 301
 QY 753 TCTTAAACCAACCTTCTGCTGTTGGAGAGACTTCATGGGAGCTGATAGATC 812
 Db 302 TCTTAAAGCCACTTACCTGCTGTTGGAGAGACTTCATGAAGACTGATGATGAC 361
 QY 813 CAGCGATACCTGTTGTTACTGAGTTCTTGGAGAGGCTGAGAAATCTGCTGTAATA 872
 Db 362 CTGCTATTCCTATAGTCACTGTTTCTTGGAGAGGATGTAATCATGATGCTGTACCA 421
 QY 873 CTTTAAAGCGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 932
 Db 422 CTTTAAAGCGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 481
 QY 933 GAGAAATGAGGATGAGAAAGATGTTGATGTTGATGATGATGATGATGATGATGATG 992
 Db 482 GAGAAATGAGGATGAGAAAGATGTTGATGTTGATGATGATGATGATGATGATGATG 541
 QY 993 CACATCAAAAGACTGTTCACTTAACTTGAAGAGCCAGAGAACTTGTATTTTG 1052
 Db 542 CAATTCCTATACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601
 QY 1053 GTCCTAGGTTTGGATCCACATCGATGAGACTGCTGAGAGAGTATTTCCAGTTA 1112
 Db 602 GTCCACAGGTTTGGATCCACATCGATGAGACTGCTGAGAGAGTATTTCCAGTTA 661
 QY 1113 GGGTAAAGATTCATCAACCTTAAGCTCCAGGACCCCTTATTCAGACAAAGAGACA 1172
 Db 662 GAGTAAAGATTCATCAACCTTAAGCTCCAGGACCCCTTATTCAGACAAAGAGACA 721
 QY 1173 TGGATTAAGGCTGGTGTACTACTAGATGATGCTCAAGCAATGCTCAATGTTTG 1232
 Db 722 TGGACAAAGG-----CATATGACTGAGATGTTGTAATTTTACATGCTGCG 775
 QY 1233 ACATTTGAGCACTCGGATGCTTGTGCTAGTATGCTTTCGCAAGGGATTCAGGATTT 1292
 Db 776 ATATAGTGAAGCAAGAGATGCTTGGCAATACGGGTTTTCAGCAAGGATTTTCAATAT 835
 QY 1293 GCATTTTGAAGTCTATGATATCTGTTGATGTTGTTGTTGTTGTTGTTGTTGTTG 1352
 Db 836 -----TTGAAGATTTGGGATCTCTGTTGATGCTGTTGCTGTTGCTGTTGCTG 889
 QY 1353 CTGTGTCACTTGAATCCATCAAGATCTGAGTGAAGGAACTGATACAGAGGCAAGTAAAC 1412
 Db 890 CGTTGACTGATGATCAAAATCTATGAGTCTGGAATGATTCAGAG-----GAAC 943
 QY 1413 TTGAACATGATGATGAGAGCTTGAAGAAATGAGAAATGTTGCTGTTGCTGTTGCTG 1472
 Db 944 TTGAACATGATGATGAGAGCTTGAAGAAATGAGAAATGTTGCTGTTGCTGTTGCTG 1532
 QY 1473 CGATTAATTTCACTTATCGAAATGTTGAGCAATCTGCTCTGTAATTAAGAAAGAGGAGAC 1592
 Db 1004 CGATCAATCTCACTGATAGGAAATGTTGAGCAATCTGCTCTGTAATTAAGAAAGAGGATTC 1652
 QY 1593 GTGTGCTGAGGAAATGTTGAGGCTTAAAGTCAATGATGCTGCAAGAGAGGCTGCAAGGTTA 1712
 Db 1064 ACCTTTCTGCAAGAAATGTTGTTATGTTCCAGATGATTCACAGAGGAGGCTGCAAGGTTA 1872
 QY 1712 ACATTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1872
 Db 1124 ACATTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1982

QY 1653 CGTTCTTGAAGACGATGCTCTATCAACAGTGCAGAGGAGAACTTACTGTTGGCTGAT 1712
 Db 1184 CGTTCTTGAAGACGATGCTCTTCTTGTGAGAGTTAGAGGAGCTGAAGTTTCCCGAAGCGAG 1243
 QY 1713 C 1713
 Db 1244 C 1244

RESULT 11
 US-10-425-114-8335
 ; Sequence 8335, Application US/10425114
 ; Publication No. US2004003488A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 8335
 ; LENGTH: 1361
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700747979_FLI
 ; US-10-425-114-8335

Query Match
 Best Local Similarity 70.8%; DB 12; Length 1361;
 Matches 872; Conservative 0; Mismatches 339; Indels 21; Gaps 4;

QY 444 GTGAGATTAATCATGTTTCTGAATTCGAAGAGGAGATTAATGCTCAAAAGCTCATATCA 503
 Db 1 GTGGGTACTAAATGTGCAAGTATGAGAGCTTCTCTTATTAAGATCTGATCTTA 60
 QY 504 AACAGGTGATGAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 560
 Db 61 GAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120
 QY 561 AACAACTATTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 620
 Db 121 AACAACTATTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
 QY 621 TTGTTCAATTTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 680
 Db 181 TAGTCTCTTTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
 QY 681 GTGTCAAGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 740
 Db 241 GTGTCAAGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
 QY 741 GTATGCGGATTTCTTAAGCAACTATCTGCTGTTGGAAGAGACTCATGGGACT 800
 Db 301 CAATGCGGATTTCTTAAGCAACTATCTGCTGTTGGAAGAGACTCATGGGACT 360
 QY 801 GATACAGATCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 860
 Db 361 GATACAGATCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
 QY 861 GTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 920
 Db 421 GTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
 QY 921 CTTTGAAGCTGAGAAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 980
 Db 481 CACTAGGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540

QY 981 CAATATCTACCCACATGCAAGACTGTTCCATCTACTTACATTTTGAAGGCCACAGAAC 1040
Db 541 CAATATATATCCCAAGGGAACTGTTCTTATTTGACATTTTACGAGCTGCTGAC 600
QY 1041 TTCTTATTTTGGTGTCTAGTTTTCATCCCAATCCATGAGACCTCTGAGAGAGTG 1100
Db 601 TACGTAATTTGGTGTCTAGTTTTCATCCCAATCCATGAGACCTCTGAGAGAGTG 660
QY 1101 ATATCCAGTTAGGGTTAAGAAATTCATACACCTTAAAGCTCCAGGACCTTTATTA 1160
Db 661 ATATTCCTGTAGGGTTAAGAAATTCATACACCTTAAAGCTCCAGGACCTCTGATCA 720
QY 1161 GACAAAGACATGATTAAGGCTGTGTCTACTAGCTAGCTAGTGTCTCAAGTCAATG 1220
Db 721 AGGCAAGATATAGACCA-----GGAGTGTAAACAAGCAATGTTTGAAGCGTAATG 774
QY 1221 TCACTATGTGACATTTGAGCACTCGGATGCTTGTGATGTTTCTGGCAAGG 1280
Db 775 TGACCATGTTGATATAGTATGATCTGATGCTGATGCTGATGCTGATGCTGATG 834
QY 1281 TATCAGGATTTGCTATATTAAGATCTATGATATCTGTGATTTGTGTCTACAGTG 1340
Db 835 TGT-----TTTCATCTTTGAAGAGTTAGGCAATACAGTTGATGTTGATGCTACAGTG 888
QY 1341 AAGTTAGTTTCTGTGTCTGCTGCTGATCCATCAAGATCTGAGTGGGAACTGATACAG 1400
Db 889 AAGTCAATGTTTCTGTGATCTGATCCATCAAGATCTGAGTGGGAACTGATACAG 948
QY 1401 AGGCAAGTGAATCTGATCCATGATGTTGAAGAGCTTGAAGAAATGCAATTTGCTTAC 1460
Db 949 AG-----GAAGTGAACATGTTGTTGAAGAACTGAGAAATCCCTGGTGGTAATCTCC 1002
QY 1461 TTGACGAGAGGGGATTAATTTGCTTATGAGAAATGAGCAATGCTGATGATAG 1520
Db 1003 TGCAATATATATCAATCTATCTCTATGAGAAATTTGAGAGATATATATATTTGG 1062
QY 1521 AAAAGACGGAGCTGTCTGAGGAAAGTGGGGTTAATGTTGATGATCTTGGCAAGAG 1580
Db 1063 AGAAGGCTTCCGTGTTCTTGAACCTTGGCATACGGGTGCAAAATGATCTCTCAAGG 1122
QY 1581 CGTCAAGGTTAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1640
Db 1123 CATCTAAGTGAAATCTCTATGTTGTAATGATGATGATGATGATGATGATGATGATG 1182
QY 1641 CCCTTCATCAGGCGTTCTTTGAAGAGATGTC 1672
Db 1183 CTCTCCACTTACGCTTCTTTGAGAGAGAGCTC 1214

RESULT 12
US-09-890-813-1
; Sequence 1, Application US/09890813
; Publication No. US20020183486A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Aspartate Kinase
; FILE REFERENCE: BB1430 PCT
; CURRENT APPLICATION NUMBER: US/09/890, 813
; PRIORITY FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/172944
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 565
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURES:
; NAME/KEY: unsure
; LOCATION: (127)
US-09-890-813-1

Query Match 28.0%; Score 555; DB 9; Length 565;
Best Local Similarity 98.9%; Fred. No. 2.3e-163;
Matches 558; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1418 CATGATGTTGAAGAGCTTGAAGAAATGCAATTTGCTTACTTCAAGAGGCGATA 1477
Db 2 CACGAGTTGAAGAGCTTGAAGAAATGCAATTTGCTTACTTCAAGAGGCGATA 61
QY 1478 ATTTCACTTATGGAATGTTGAGCAATGCTCTGATATCTAGAAAGAGCGGCTG 1537
Db 62 ATTTCACTTATGGAATGTTGAGCAATGCTCTGATATCTAGAAAGAGCGGCTG 121
QY 1538 CTGAGAAAGTGGGGTTAATGTTGATGATCTGCAAGAGCGTCAAGGTTAATG 1597
Db 122 CTGAGAAAGTGGGGTTAATGTTGATGATCTGCAAGAGCGTCAAGGTTAATG 181
QY 1598 TCGTGAATATGTCATGATAGCGATGCAAGAGCGTCAAGAGCGTCAAGGTTA 1657
Db 182 TCGTGAATATGTCATGATAGCGATGCAAGAGCGTCAAGAGCGTCAAGGTTA 241
QY 1658 TTTGAAGAGATGTCCTATCAAGAGCGTCAAGAGCGTCAAGAGCGTCAAGG 1717
Db 242 TTTGAAGAGATGTCCTATCAAGAGCGTCAAGAGCGTCAAGAGCGTCAAGG 301
QY 1718 TAGGCTTTGCTGGTCCAGGCGTCTTATCTGTTATGATTCCTGCTCCATGAG 1777
Db 302 TAGGCTTTGCTGGTCCAGGCGTCTTATCTGTTATGATTCCTGCTCCATGAG 361
QY 1778 GATGGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1837
Db 362 GATGGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 421
QY 1838 AGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1897
Db 422 AGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 481
QY 1898 CGAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1957
Db 482 CGAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 541
QY 1958 AT 1981
Db 542 AT 565

RESULT 13
US-10-424-599-28060
; Sequence 28060, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yinhua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424, 599
; PRIORITY FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 28060
; LENGTH: 2490
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURES:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_125339C.1
US-10-424-599-28060

Query Match 26.9%; Score 533.2; DB 12; Length 2490;
Best Local Similarity 63.9%; Fred. No. 5.1e-156;
Matches 987; Conservative 0; Mismatches 384; Indels 174; Gaps 5;

QY 284 GTGATGAAGTTCGGGGGCTCTCGGTGCTGCGCGCGAGAGATGCTGAGGTGGCGGC 343

```

Db      607 GTCAAGAACTTGGTGGCTCTCCGCTGCAATGCTAAAGAAAGAGAGGTTGGCAAC 666
Qy      344 CTCATCTGACCTTCCCGAGAGAGCGCCCGTGGTCTCTCTGCGCAATGGGAAAACC 403
Db      667 CTATATCTGAGCTTCCCGAAGAGGCAATTAATGTTCTCTGCGCATGGGAAAAGACA 726
Qy      404 ACCAACAACCTTCTCTGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 463
Db      727 ACTAACAATGCTTTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 786
Qy      464 GAAATCGAAGAGTGAATATGCTCAAAAGCCTACATCAAGAGGAGTGAAGAGAGAGAGAG 523
Db      787 AGTATTTGATGAGCTCAACATATATAAAGATCTCATCTCAGAGCTGTGGAACAGCTTGA 846
Qy      524 CTTCAGAGNA---TCTGTATATACAAGCTTTATGAGCTGAGCAACTATGAAAGGATC 580
Db      847 GTGACAGAAATGTTATTTAGAGAGATCTTAAGAAATTTGAGCACTTTCTTAAGGGGATA 906
Qy      581 GCTATGATGAAGAGAGCTGAGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 640
Db      907 GCTATGATGAAGAGAGTGAAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 966
Qy      641 AGTCCACAGAGATTTTCTGCTTATTTGAACAAATGCTGTCAAGAGAGAGAGAGAGAGAT 700
Db      967 AGTCCACAGAGATTTTCTGCTTATTTGAACAAATGAGAGAGAGAGAGAGAGAGAGAT 1026
Qy      701 GAGCATTGATATTTGTTTCAATTAACAAGATGATGATGATGATGATGATGATGATGAT 760
Db      1027 GATGATTTGAGATGAGATTAATATACATGATGATGATGATGATGATGATGATGAT 1086
Qy      761 GCAACCTATCCCTGCTGCTGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 820
Db      1087 GCAACATATCCCTGCTGCTGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1146
Qy      821 CCTGTTGTTACTGAGGCTCTGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 880
Db      1147 CCAATTTGTTACAGGCTCTGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1206
Qy      881 CAGAGTGTGATGATCTGATCTGATCAACATGATGATGATGATGATGATGATGATGAT 940
Db      1207 AAGAGTGAAGAGATTAATCTGATCAACATGATGATGATGATGATGATGATGATGAT 1266
Qy      941 CAGATGATGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1000
Db      1267 CAGATGATGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1326
Qy      1001 AAGACTGTTCACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 1060
Db      1327 AAACCTGTTCACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 1386
Qy      1061 GTTT----- 1064
Db      1387 TGTTAACACAACTTATTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1446
Qy      1065 ----- 1064
Db      1447 AATAGAGAACTGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1506
Qy      1065 -----TGACATCAACATGATGAGAGCT 1087
Db      1507 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1566
Qy      1088 GCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1147
Db      1567 GCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1626
Qy      1148 ACCCTATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1207
Db      1627 ACTCTATGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Qy      1208 CTCACGTAATGTCATATGTTGACATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1267

```

```

Db      1681 TTGAAGATATGACCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
Qy      1268 TTTCTGGCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1327
Db      1741 TTTCTGGTAAGAT-----TTTCAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1794
Qy      1328 GTTCTCAACAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1387
Db      1795 GTACTCAAGATGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1854
Qy      1388 GAATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1447
Db      1855 GAGCTAATTCACAG-----GAACTGATGATGATGATGATGATGATGATGATGATGATGAT 1908
Qy      1448 ATGTTGCTCTACTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1507
Db      1909 GTGGAATATCTTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1968
Qy      1508 TCTCTGATCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1567
Db      1969 TCCTAATATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2028
Qy      1568 ATCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1627
Db      2029 ATCTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2088
Qy      1628 GCACTCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1672
Db      2089 CAGTGTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2133

RESULT 14
US-10-425-114-7821
; Sequence 7821, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7821
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700684751_FLI
US-10-425-114-7821

Query Match      25.9%; Score 515; DB 12; Length 1671;
Best Local Similarity 63.8%; Pred. No. 2e-150;
Matches 961; Conservative 0; Mismatches 371; Indels 174; Gaps 5;

```


Db 184 AGACACTGTGAAACAGTTGGAGTGGACAGAAATGTTATTTGAGAACATCTTGAAGAAATG 243
 QY 560 GAGCAACTATTGAAAGATATCCCTATGATGAAAGAGCTGACGCTTAGAACACAGTACTAC 619
 Db 244 GAGCAACTTCTAAAGGGGATAGCTATGAGAAAGGTTGACTCCACGAGCTCAAGACTAT 303
 QY 620 CTGTGTTCAATTTGAGAAATGACATGTCACAGAGATTTTTCGCTTAATTTGACAAAT 679
 Db 304 TTAGTTTCAATTTGAGAGTGAATGTCACACTAGAGATTTTGCCTGATCTTAACACATTA 363
 QY 680 CGTGTCAAGGACGAGATGATACGCAATTTGATTTGGTTTCACTTACACTGATGAATTT 739
 Db 364 GAGGTTAAGGCGGCGCAATATGATGCAATTTGAGATGGGTATTTATTAACAAGTGAAGCTTC 423
 QY 740 GGTATGCGGATATCTTAGAACAACTATCTGCTGTGGGAAGAGACTCATGGGAC 799
 Db 424 ACAATGCTGACATTTTGAAGCAACATATCTGCTGTGGCAAGAGTTACATAGTAT 483
 QY 800 TGGATACAGGATCCAGCATACCTGTGTTACTGGGTTCTTGGAGGSCCTGGAATCT 859
 Db 484 TGGGTTCTGATCTGCTCTATTCOAATTTGTACAGGCTTCTTGGAAAGCAGAAATCA 543
 QY 860 GGTGCTGTACTACTTTAGCCGAGGTGGTAGTGACTTGAAGCTACAAACCATTTGATTA 919
 Db 544 TGTGCGGTACACATTTGGGTAGAGTGAAGTGAATTTTAACTGCTACAAACATTTGGGAAA 603
 QY 920 GCCTTGGGACTGAGAGAAATTCAGGTATGAGAAAGATGTTGATGTTACTTACTTGTGAT 979
 Db 604 GCGCTAGGGTTGCGAGAAATCCAGGATGGAAGATGTTGATGGGTCTTAACCTGTGAT 663
 QY 980 CCAAAATCTACCCAGATGCCAAGACTGTTCCATCTTAACCTTGAAGAGCCACAGAA 1039
 Db 664 CCAAAATATGCCCAAGCAAAACCTGTTCCATTTGACATTTGATGAAGGCTGCTGAA 723
 QY 1040 CTTCCTTATTTTGGTCTCAG----- 1060
 Db 724 CTTCGATACCTTGGTCTCAGTGTAAACACATATTTGATAGGAGAACTCAATACTG 783
 QY 1061 ----- 1060
 Db 784 CAATGTACACTGGGAAAGAAATAGAGAACTGAAGAGGCTACAGGCTTAATTC 843
 QY 1061 -----GTTTGG 1066
 Db 844 AAAAGGCTATCTCTTGACCATGAGAGAACCGGATAGAAATTCACAAAGTCTCTC 903
 QY 1067 CATCCAAATCGATGAGACCTGTGTAGAGAAAGTGATATTTCCAGTTAGGGTTAAGAAATTC 1126
 Db 904 CATCCAGCTCTATGAGACCGGCTAGGAAAGTGAATTTCCGTTAGGGTTAAAAATTC 963
 QY 1127 TACAACCTTAAAGCTCCAGGACCCCTTATTTACCAAGAAAGAGACATGATTAANGTCTG 1186
 Db 964 TACAACCTTAAAGCTCCAGGACTCTGATTTGCTAAGACAGAGGAAATAGCA-----G 1017
 QY 1187 GTTGTACTAATGATGCTCAAGTCAATGTCATGTTGATGTCATTTGATGACACT 1246
 Db 1018 GCATTTATTAACAAGCAATGTTTGAACCTAATGTGACCATGTTGATGATTTGTAGCACT 1077
 QY 1247 CGATGCTTGGTATGATGTTTTCGCAAGGATATGAGTATTTGATTAATGAGAT 1306
 Db 1078 CGATGCTTGGTATGATGTTTTCGCTGTAAGAT-----TTCAATATTTGMAAG 1131
 QY 1307 CTATGATATCTGTGATTTGTGCTACAGTGAAGTATGTTTGTGTCTCACTTAT 1366
 Db 1132 TTAGGATATCACTTATGTTAGCTACAGTGAAGTATTTCTTGAACATTTGAT 1191
 QY 1367 CCATCAAGATCTGAGTGGGAACTGATACAGGAGGAGGAGTGAACCTTGAACATGAT 1426
 Db 1192 CCATCAAGATCTGAGGAGAGAGCTAATTTCAAG-----GAACCTTGAATGTTCTC 1245
 QY 1427 GAAGACTTGAGAAATAGCAATGTTGCTCTACTTCAGAGAGGCGATTAATTCCTT 1486

Db 1246 GAAGAACTGGAAAAATTTGCTGTGTAAATCTCTTAAAGACAGATCATTAATCTCTC 1305
 QY 1487 ATCGGAAATGTGAGCAATCGTCTGTATCTAGAAAAAGACGGGACGTGTGAGAGAAA 1546
 Db 1306 ATTTGAAATGTTCAAGAGTCACTAATATTTGAGAAAGGCTTTTCAATGTTCTTCAACT 1365
 QY 1547 ACTGGGTTAAATTTGATGATCTCGAAGAGGCTCAAGGTTAAACATGTGCTGATA 1606
 Db 1366 CTGGGATCACTGTTCAAATGATCTCAGGGTGCACTTAAGTGAATATCTGTAGTT 1425
 QY 1607 GTCCATGATPAGCATGCAAAAGCACTGTGAAAGCCCTTCATCAGGCGTCTTTGAAGAC 1666
 Db 1426 ATAAATGACATGAAAGACCAAGTGTCTCAGGCTCTCCAAAGGCTTTTGAAGAC 1485
 QY 1667 GATGTC 1672
 Db 1486 GAATCTC 1491
 RESULT 15
 US-10-425-114-7451
 ; Sequence 7451, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Tabaska, Jack E.
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 7451
 ; LENGTH: 1260
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700663367_FLI
 US-10-425-114-7451
 Query Match 23.7%; Score 470.4; DB 12; Length 1260;
 Best Local Similarity 71.6%; Pred. No. 1.8e-136;
 Matches 666; Conservative 0; Mismatches 246; Indels 18; Gaps 3;

QY 743 AATGGGATATTTTGAAGCAACCTATCTCTGTTGCGAAGACTTCAATGGGACTGG 802
 Db 3 AATGGGACATTTTGAAGCAACTATCTCGGACAGTCCCAAGATTAATGAGATGG 62
 QY 803 ATACAGGATCCAGCATACCTGTTTACTGCGTTCTTGGGAAGGCTGAAATCTGAT 862
 Db 63 CTCTGATCTCGCAATGCAATTTCTTACAGGCTTCTTGGAAAGGCCAGAAATCATGT 122
 QY 863 GCTGTACTACTTTAGSCCGAGGTGTAGTGAATGATGATGCTGACACCATTTGTAAGCC 922
 Db 123 GAGGTGACAACTGAGTGAAGGAGGAGGAGTATTTGACAGCTACAGCAATTTGGGAAAGCA 1082
 QY 923 TTGGGACGAGGAATTCAGTATGGAAGAGTGTAGTGTACTTACTTGTGATCCA 982
 Db 183 CTAGGTTTACCTGAGATCCAGATGGAAGATGTTGATGTTGCTTCACTGTGATCCA 242
 QY 983 AATATTCACCAATGCAAGAACTGTTGCATTAATTTAAACATTTGAAGGCCACAGAACTT 1042
 Db 243 AATATATACCCAAAGCCGAAACCTGTTCTTAATTTGACATTTGACGAGGCTGTGAACCTA 302
 QY 1043 GCTTAATTTTGGTGTCAAGTTTGTGATCAACATGATGAGACTGTGTAGAAAGTGAT 1102
 Db 303 GGTACTTTGGTGTCAAGTTTGTGATCAACATGATGAGACTGTGTAGAAAGTGAT 362
 QY 1103 ATTCAGTTTGGTTAGAAATTCATACAAACCTTAAGCTCCAGGACCTTATTACAGA 1162

Wed Mar 24 14:12:54 2004

us-09-890-813-5.rnpb

Page 14

Db	363	ATTCCGTTAGGGTTAAATAATTTCTTACATCTTAAAGTCTCAGATCTCTCATACCAAG	422
Oy	1163	CAAGAGACATGGATTAANGTCTGTGGTTGTACTAACTAGTGTCAAGTCAAAATGTC	1222
Db	423	GCAAGGAAATAGGCA-----GGCAGTGTAAACAGCATTTGTTGAAGCATATGTG	476
Oy	1223	ACTATGTGACATTTGTGAGCACTCGAGTGGTGTGAGTATGTTTCTGCAAGGGTA	1282
Db	477	ACCAATTTGCAATATGTAGTACTCGAATGTTGGTCAATATGTTTCTTCTAAGGTG	536
Oy	1283	TCAGTATTTGGTATATTGAAGATCTATGATATCTGTGGATTGTGTCTACAGTAA	1342
Db	537	T-----TTTCAATCTTTGAAGATTAAGCATATCAAGTTATGTTGTACTCAAGTAA	590
Oy	1343	GTTAGTGTTCGTGTCTCACTTGATTCATCCAAAGTCTGGAGTGGGAATGATTAAGAG	1402
Db	591	GTCAGTGTTCCTTACATGTGATTCATCAAAAGCTAATGAGCAAGAGCTAATTAGAG	650
Oy	1403	GCAAGTGAATTTAGCCATGTATGTTGAAGAGCTTGAAGAAATAGCAATTTGTGTACTT	1462
Db	651	-----GAATTTGACCATGTTGTGAAGAACTCGAGAAATTCGTGTGATATCTCTG	704
Oy	1463	CAGCAGAGGGCCATTAATTTCACTTATTCGGAATGTGAGCAATCTGTTCTGATACTAGA	1522
Db	705	CAGAAATGATTCATATCTCTCTCATTTGGAAATTTTGAAGATCAATCACTAATATTGGAG	764
Oy	1523	AAGACGGACGTGTCTGTGAGGAAAAGTGGGGTTTATTTGATGATCTCGAGAGGACG	1582
Db	765	AAGGCTTCCGTGTTCTTGCAACCTTGCACTACGATGCAATGATCTCTTAGGGTCA	824
Oy	1583	TCAAAGTTAACATGTCGCTGATATGTCATGATAGCGATCAAGGCAAGCACTCGTAAAGCC	1642
Db	825	TTTAAAGTGAACATCTCAATTGGTTGTAATATACAGTGAACGAGCAAGTGTGAGGCT	884
Oy	1643	CTTCAATCAAGCGCTTCTTGAAGAGAGTGC	1672
Db	885	CTCCACTTGAACCTTCTTGAAGAGTGAAGCTC	914

Search completed: March 23, 2004, 13:14:24
Job time : 711 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 07:05:34 ; Search time 5264 Seconds

(without alignments)
11260.720 Million cell updates/sec

Title: US-09-890-813-5

Perfect score: 1985

Sequence: 1 gcacacagacagacagagag.....ataaacattcaatctt 1985

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

EST:*

1: em_estda:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vtl:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_tod:*
26: em_gss_png:*
27: em_gss_vtl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	795.2	40.1	3000	AY109464	Zea mays
2	635.4	32.0	682	CD440698	EL01N0558
3	584.2	29.4	756	CB666188	OS0NED12P
4	578	29.1	665	CD441517	EL01N0558

5	513.4	25.9	549	14	CA250542	CA250542	SCRUP111
6	477.4	24.1	512	10	AM285751	AM285751	Lcl_223-B
7	468.4	23.6	502	10	AM285752	AM285752	Lcl_223-B
8	455.4	22.9	1103	14	CK206083	FGAS01765	CK206083
9	445.2	22.4	1101	14	CK206100	FGAS01767	CK206100
10	444	22.4	827	14	CB679618	OSJNEF03F	CB679618
11	438.6	22.1	792	13	BQ506340	EST613755	BQ506340
12	437.6	22.0	764	14	CF449571	EST695916	CF449571
13	436.2	21.9	548	10	BQ653377	EST695916	BQ653377
14	434.2	21.9	727	13	BQ986067	HP16004x	BQ986067
15	431	21.7	607	13	CB623763	OS1IEA11L	CB623763
16	425.8	21.5	822	14	AI901511	61806D07	AI901511
17	420.2	21.2	433	9	CA931918	MTU47A_P2	CA931918
18	419.8	21.1	670	14	CA932103	MTU47A_P2	CA932103
19	418.8	21.1	680	14	CA932103	MTU47A_P2	CA932103
20	417.8	21.0	653	28	BZ328792	OSABE871C	BZ328792
21	417.8	21.0	839	29	CG376871	OSABE871C	CG376871
22	416.4	21.0	911	29	CG200786	P0F055TD	CG200786
23	416.4	21.0	960	28	CG413956	P0F055TD	CG413956
24	416.2	21.0	836	29	CG151954	P0F055TD	CG151954
25	412.2	20.8	692	12	BG441599	GA_Ea001	BG441599
26	407.6	20.5	733	13	BQ939285	OGF6A07_Y	BQ939285
27	402.2	20.3	690	12	BQ595665	EST494347	BQ595665
28	398.6	20.1	658	12	BU301147	BU301147	BU301147
29	395.6	19.9	689	13	CA064976	SCACAD103	CA064976
30	394.4	19.9	660	14	CD054614	H00F02x	CD054614
31	388.2	19.6	913	28	CG373518	P0F031TD	CG373518
32	383.4	19.3	891	29	CG322733	OG0F171TV	CG322733
33	373.4	18.8	895	29	CG322743	OG0F171TV	CG322743
34	366.2	18.4	605	14	CF093872	QXN16C18	CF093872
35	364.8	18.4	697	14	CA219909	SCSFL403	CA219909
36	364.6	18.4	559	14	CF098434	QNR7E22_Y	CF098434
37	364	18.3	706	14	CF395448	RTD52_11	CF395448
38	360.4	18.2	615	14	CF322694	HDN_-01-M	CF322694
39	357.6	18.0	645	10	BE249753	NF021H03L	BE249753
40	356.6	18.0	691	13	BQ148318	NF05B909F	BQ148318
41	356.2	17.9	452	12	BG158155	RH1Z2_28	BG158155
42	356	17.9	356	9	A1629668	486104E06	A1629668
43	355.6	17.9	755	13	BQ005561	OGG8H17_Y	BQ005561
44	352	17.7	774	14	CB619049	OS1IEA03F	CB619049
45	349	17.6	618	10	BE458252	EST413592	BE458252

ALIGNMENTS

RESULT 1	AY109464	3000 bp	mrna	linear	HTC 17-OCT-2002
LOCUS	AY109464				
DEFINITION	Zea mays CL754_1 mRNA sequence.				
ACCESSION	AY109464				
VERSION	AY109464.1	GI:21213193			
KEYWORDS	HTC.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD				
AUTHORS	1 (bases 1 to 3000)				
TITLE	Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitesitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.				
JOURNAL	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes				
REFERENCE	Unpublished (2002)				
AUTHORS	2 (bases 1 to 3000)				
TITLE	Coe, E.H.				
JOURNAL	Direct Submission				
COMMENT	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA				
	If you are interested in getting corresponding physical clones, these are publicly available from ZMDB and may be found by BLAST searching at msl.maizemap.org , zmdb.www.zmdb.iastate.edu , TIGR, www.tigr.org , or NCBI, www.ncbi.nlm.nih.gov . When the source of the				

QY	1200	GCATGATGCTCAAGCAACAAATGCTACTAATGTTGACATTTGGACACTCGGATGCTTGCTC	1259
Db	686	GCATGATGCTCAAGCAACAAATGCTACTAATGTTGACATTTGATATGTAAGCACTCGATGCTTGCTC	745
QY	1260	AGTATGCTTTT	1270
Db	746	AATTTGCTTTT	756
RESULT 4	CD441517/c	665 bp mRNA linear EST 03-JUN-2003	
LOCUS	EL01N0558F11.9	EndospERM_5 Zea mays cDNA, mRNA sequence.	
DEFINITION	CD441517		
ACCESSION	CD441517.1	GI:31357160	
VERSION	EST		
KEYWORDS	Zea mays		
SOURCE	Zea mays		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
REFERENCE	1 (bases 1 to 665)		
AUTHORS	Lai,J., Dey,N., Kim,C.S., Becraft,P., Larkins,B., Linton,E. and Messing,J.		
TITLE	Sequencing of the maize endospERM ESTs		
JOURNAL	Unpublished (2002)		
COMMENT	Contact: Lai, Jinsheng Dr. Joachim Messing's lab Waksman Institute, Rutgers University 190 Frelinghuysen Rd., Piscataway, NJ 08854, USA Tel: 732-445-1801 Fax: 732-445-5735 Email: jlai@waksman.rutgers.edu Seq primer: 77.		
FEATURES	Location/Qualifiers		
source	1..665		
	/organism="Zea mays"		
	/mol_type="mRNA"		
	/cul_e1va="w22"		
	/db_xref="taxon:4577"		
	/c1eise_type="EndospERM of 7-33DAP"		
	/clone_id="EndospERM_5"		
	/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"		
ORIGIN			
Query Match	29.1%; Score 578; DB 14; Length 665;		
Best Local Similarity	94.1%; Pred. No. 6.1e-112;		
Matches 637; Conservative 0; Mismatches 26; Indels 14; Gaps 3;			
QY	1166	AGAGACATGGATNANGCTGGTGTACTACTAGCATAGTGCATCAATGATGCACT	1228
Db	665	AGACAAAGACATGGATAGGTTTATCTATCTAGCATAGTGCATCAATGATGCACT	606
QY	1226	ATGTTGACATTTGTGACACTCTGGATGCTTGGTCAGATAGGTTTCTGGCAGGGTATCA	1288
Db	605	ATGTTGACATTTGTGACACTCTGGATGCTTGGTCAGATAGGTTTCTGGCAGGGTATTT	546
QY	1286	GGATTTTGCATATATGAAGATCTATGATATCTGATGATGTTGTTGTCACAGTGAAGTT	1345
Db	545	GCTATAT-----TTGAATCTATGATATATCCGTGATGTTGTTGTCACAGTGAAGTT	492
QY	1346	AGTGTTCCTGTGCTACTTGATTCATCAAAAGATCTGGAGTAGGGAACCTGATACAGCGCA	1408
Db	491	AGTGTTCCTGTGCTACTTGATTCATCAAAAGATCTGGAGTAGGGAACCTGATACAGCG--	435
QY	1406	AGTGAACCTTACCATGATAGTGAAGAGCTTGAGAAATATAGCAATGTTGTTCTACTTCAG	1465
Db	434	---GAACTTACCATGATGATGAAGAGCTTGAGAAATATAGCAATGTTGTTCTACTTCAG	378
QY	1466	CAGAGGCGCATATTTCACTTATCGGAAATGTGAGCAATGCTCTGTGATACTAGAAAAG	1522

[illegible]

source of RNA and library construction can be obtained at
http://succest.lad.ic.unimcamp.br/public"

ORIGIN

Query Match 25.9%; Score 513.4; DB 14; Length 549;
Best Local Similarity 96.0%; Pred. No. 5.7e-116;
Matches 526; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 576 GATGCGCTATGATGAAGAGCTGACGCCCTAGAGACAGACCTGTTTCACTTTGAG 635
DB 1 GATGCGCTATGATGAAGAGCTGACGCCCTAGAGACAGACCTGTTTCACTTTGAG 60
QY 636 AATGCATGTCACACAGATTTTCTGCTTATTTGAACAAATTCGTCAAGCCAGCC 695
DB 61 AATGCATGTCACACAGATTTTCTGCTTATTTGAACAAATTCGTCAAGCCAGCC 120
QY 696 AGATATACCATTTGATATTTGTTTCACTACACAGATTTGTTATGCGATATCT 755
DB 121 AGATATACCATTTGATATTTGTTTCACTACACAGATTTGTTATGCGATATCT 180
QY 756 TAGAAGCAACTATCCTGCTGTTGCGAAGAGACTTCAATGGGACTGATACAGATCCAG 815
DB 181 TAGAAGCAACTATCCTGCTGTTGCGAAGAGACTTCAATGGGACTGATACAGATCCAG 240
QY 816 CGATACCTGTTGTTACTGCGTTCTTTGGGAGAGGCTGGAATCTGCTGTTACTT 875
DB 241 CGATACCTGTTGTTACTGCGTTCTTTGGGAGAGGATGGAATCTGCTGTTACTT 300
QY 876 TAGGCGAGGTGATGATGACTGACTGCTACCAACCATTTGTTAAAGCTTGGAGACTGAG 935
DB 301 TAGGCGAGGTGATGATGACTGACTGCTACCAACCATTTGTTAAAGCTTGGAGACTGAG 360
QY 936 AAATTCAGGTATGGAAGAGATGTTGATGTTACTTCTTGTATCCAAATTTCTAACCCAC 995
DB 361 AAATTCAGGTATGGAAGAGATGTTGATGTTACTTCTTGTATCCAAATTTCTAACCCAC 420
QY 996 ATCCAAAGACTGTTCCATCTTACATTTGAAGAGGCCACAGAACTTCTATTTTGGTG 1055
DB 421 ATCCAAAGACTGTTCCATCTTACATTTGAAGAGGCCACAGAACTTCTATTTTGGTG 480
QY 1056 CTCAGGTTTGCATCCACATGATGAGACTGCTTGAAGAGGTGATATTCAGTTAGGG 1115
DB 481 CTCAGGTTTGCATCCACATGATGAGACTGCTTGAAGAGGTGATATTCAGTTAGGG 540
QY 1116 TTAAGAT 1123
DB 541 TTAAGAT 548

RESULT 6
AM285751 512 bp mRNA linear EST 19-JUL-2000
LOCUS UGI.223.B06.b1_A002 Light Grown 1 (LGI) Sorghum bicolor cDNA, mRNA
DEFINITION
sequence.
ACCESSION AM285751
VERSION AM285751.2 GI:6858218
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Sorghum.
1 (bases 1 to 512)
Cordonier-Pratt, M.-M., Gingle, A., Marsala, C. and Pratt, L.H.
An EST database from Sorghum: light-grown seedlings
Unpublished (2000)
On Jan 6, 2000 this sequence version replaced gi:6675595.
Contact: Cordonier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210

Email: mmpratt@uga.edu
Sequences have been trimmed to exclude polyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 506
POLYA=No.
FEATURES
source
location/Qualifiers
1..512
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone_lib="Light Grown 1 (LGI)"
/note="Organ: 10- to 14-day-old light-grown (greenhouse)
seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2:
EcoRI; The library was made from poly-A RNA in the cloning
vector lambda Zap II. Clones to be sequenced were
prepared by mass excision."

ORIGIN

Query Match 24.1%; Score 477.4; DB 10; Length 512;
Best Local Similarity 96.5%; Pred. No. 4.6e-107;
Matches 500; Conservative 0; Mismatches 12; Indels 6; Gaps 1;

QY 751 TATCTTAGAGCACTTCTGCTGTTGCCAAGAGACTTCATGGGACTGATACAGGA 810
DB 1 TATCTTAGAGCACTTCTGCTGTTGCCAAGAGACTTCATGGGACTGATACAGGA 60
QY 811 TCCAGGATACCTGTTGTTACTGCTGTTCTTTGGGAGGGCTGGAATCTGCTGTAA 870
DB 61 TCCAGGATACCTGTTGTTACTGCTGTTCTTTGGGAGGGCTGGAATCTGCTGTAA 120
QY 871 TACTTAGCCGAGGTGATGATGACTGCTACCAACCATTTGTTAAAGCTTGGAGCT 930
DB 121 TACTTAGCCGAGGTGATGATGACTGCTACCAACCATTTGTTAAAGCTTGGAGCT 180
QY 931 GAGAGAAATTCAGTATGGAAGAGATGTTGATGTTACTTCTTGATCCAAATATCTA 990
DB 181 GAGAGAAATTCAGTATGGAAGAGATGTTGATGTTACTTCTTGATCCAAATATCTA 240
QY 991 CCCACATGGAAGACTGTTCCATCTTACATTTGAAGAGGCCACAGAACTTGTATTT 1050
DB 241 CCCACATGGAAGACTGTTCCATCTTACATTTGAAGAGGCCACAGAACTTGTATTT 300
QY 1051 TGGTGTCAAGTTTGCATCCACATGATGAGACTGCTTGAAGAGGTGATATTCAGT 1110
DB 301 TGGTGTCAAGTTTGCATCCACATGATGAGACTGCTTGAAGAGGTGATATTCAGT 360
QY 1111 TAGGTTAAGATTCATACAACTTAAGCTCCAGGACCTTTATTCAGACAAAGGA 1170
DB 361 TAGGTTAAGATTCATACAACTTAAGCTCCAGGACCTTTATTCAGACAAAGGA 420
QY 1171 CATGATTAAGTCTGTTGTTACTACTAGATAGTGTCAAGTCAATGTCACTATGTT 1230
DB 421 CATGATTA-----GTTGTTACTACTAGATAGTGTCAAGTCAATGTCACTATGTT 474
QY 1231 GGCATTTGAGCACTCGGATGCTTGTGTCAGTATGTT 1268
DB 475 GGCATTTGAGCACTCGGATGCTTGTGTCAGTATGTT 512

RESULT 7
AM285752 502 bp mRNA linear EST 19-JUL-2000
LOCUS UGI.223.B07.b1_A002 Light Grown 1 (LGI) Sorghum bicolor cDNA, mRNA
DEFINITION
sequence.
ACCESSION AM285752
VERSION AM285752.2 GI:6858219
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 502)
Cordonnier-Pratt, M.-M., Gingle, A., Marsala, C. and Pratt, L.H.
An EST database from Sorghum: light-grown seedlings
Unpublished (2000)
On Jan 6, 2000 this sequence version replaced gi:675596.
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmp@prattuga.edu

Sequences have been trimmed to exclude polyA, vector and regions below phred quality 16. The threshold for highest quality sequence is 20.
Seq primer: JEN REV
High quality sequence stop: 353
POLYA=No.
Location/Qualifiers
1..502
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone_lib="Light Grown 1 (LGI)"
/note="Organ: 10- to 14-day-old light-grown (greenhouse) seedlings; Vector: Lambda Zap; Site 1: XhoI; Site 2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."

ORIGIN

Query Match 23.6%; Score 468.4; DB 10; Length 502;
Best Local Similarity 96.7%; Pred. No. 7.8e-105;
Matches 491; Conservative 0; Mismatches 11; Indels 6; Gaps 1;

QY 751 TATCTTNGAAGCACTATTCCTGCTTTCGGAAGACCTTCAATGGGAGCTGGATACAGA 810
Db 1 TATCTTNGAAGCACTATTCCTGCTTTCGGAAGACCTTCAATGGGAGCTGGATACAGA 60
QY 811 TCCAGCGATACCTGTTTACTGGGTTCCCTTGGGAGGCTGGAATCTGTGCTGTAAAC 870
Db 61 TCCAGCGATACCTGTTTACTGGGTTCCCTTGGGAGGCTGGAATCTGTGCTGTAAAC 120
QY 871 TACTTTAGCGCGAGGTGTAGTGAATCTGCTGACACCAATGGTAAAGCTTGGGACT 930
Db 121 TACTTTAGCGCGAGGTGTAGTGAATCTGCTGACACCAATGGTAAAGCTTGGGACT 180
QY 931 GAGAGAAATCAGGTGTGAAAGATGTGATGCTTACTTGTGATCCAAATATCTA 990
Db 181 GAGAGAAATCAGGTGTGAAAGATGTGATGCTTACTTGTGATCCAAATATCTA 240
QY 991 CCCACATGCAAAAGCTGTTCCATCTTAACTTGAAGGCCACAGAACTGCTTATTT 1050
Db 241 CCCACATGCAAAAGCTGTTCCATCTTAACTTGAAGGCCACAGAACTGCTTATTT 300
QY 1051 TGGTGCTCAGGTTTGTGATCCATCGATGAGAGCTGCTAGAGAAAGTGATATCCACT 1110
Db 301 TGGTGCTCAGGTTTGTGATCCATCGATGAGAGCTGCTAGAGAAAGTGATATCCACT 360
QY 1111 TAGGGTTAAGAAATTCATCAACCTTAAGCTCCAGGACCCCTTATTTACCAAGAAAGGA 1170
Db 361 TAGGGTTAAGAAATTCATCAACCTTAAGCTCCAGGACCCCTTATTTACCAAGAAAGGA 420
QY 1171 CATGATTAANGGTGCTGGTTTACTACTAGCATAGTGTCAAGTCAATGTACTATATT 1230
Db 421 CATGATTA-----GGTTTACTACTAGCATAGTGTCAAGTCAATGTACTATATT 474
QY 1231 GGACATTGTAGACACTCGATGCTTGGT 1258
Db 475 GGATATCGTAGCACTCGATGCTTGGT 502

RESULT 8
CK206083 1103 bp TRNA linear EST 08-DEC-2003
LOCUS
DEFINITION
FGS017659 Triticum aestivum FGS: Library 5 GATE 7 Triticum
aestivum cDNA, mRNA sequence.
ACCESSION
CK206083
VERSION
CK206083.1 GI:39568473
KEYWORDS
SOURCE
ORGANISM
Triticum aestivum (bread wheat)
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Triticum.
1 (bases 1 to 1103)

REFERENCE
AUTHORS
Allard, F., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,
Geneswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A.,
Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilson, D.,
Penhiket, C., Roach, J.L. and Sarhan, F. Muzak, I., Nilson, D.,
Functional Genomics of Abiotic Stresses in Wheat and Canola Crops
Unpublished (2003)
Contact: Wm L Crosby
Bioinformatics
University of Saskatchewan, Department of Computer Science
1101 Engineering Building, 57 Campus Drive, Saskatoon,
Saskatchewan, S7N 5A9, Canada
Tel: 306 966 1769
Fax: 306 966 2033

Email: fgas_estes@usask.ca
This sequence is the direct result of the Base calling software
phred (default parameters). It is the raw base calls. To aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [18,813].
Plate: L5B06 row: D column: 14.

FEATURES

source

1..1103

Location/Qualifiers

/organism="Triticum aestivum"

/mol_type="mRNA"

/db_xref="taxon:4565"

/clone_lib="Triticum aestivum FGS: Library 5 GATE 7"

/note="Vector: pCMV.SPORT6; Crown and developmental stages

of spike formation in wheat cultivar Norstar. 4 mRNA

populations were combined before constructing the library.

The first mRNA population is from 1cm crown sections after

30 days of cold acclimation. The second is from 1cm crown

sections after 11 days of deacclimation (before

deacclimation plants were fully vernalized for 49 days).

The third is from different developmental stages of spike

formation (5 to 50mm) that still have not emerged from the

leaf (dissection required). The last is from different

developmental stages of spike and seed formation after

having emerged from the leaf (visible). First strand

synthesis in this library was done in the presence of

methylated dCTP thereby protecting from internal cleavage

with NotI."

ORIGIN

Query Match 22.9%; Score 455.4; DB 14; Length 1103;
Best Local Similarity 73.5%; Pred. No. 2e-101;
Matches 627; Conservative 0; Mismatches 208; Indels 18; Gaps 3;

QY 846 AGGCGTGAAGATCTGGTGTGAATCTTTAGGCCGAGGTGTAGTGAATCTGACTGTA 905
Db 26 AGGCGTGAAGATCTGGTGTGAATCTTTAGGCCGAGGTGTAGTGAATCTGACTGTA 85
QY 906 CAACCATTTGGTAAAGCTTGGGAGTGAAGAAATTCAGGTATGAAAGATTTGATGATG 965
Db 86 CAACCATTTGGTAAAGCTTGGGAGTGAAGAAATTCAGGTATGAAAGATTTGATGATG 145
QY 966 TACTTAATTTGATCCAAATATCTACCAATGCAAGACCTGTTCCATCTTAACATTTG 1025
Db 146 TGTGACGTGTATCCAAATATTTATGCAAAATGGGTACAGTACCTTACTGACTTTTG 205

QY 1026 AAGAGGCCACAGAACTGGTTATTTTGGTCTCAGGTTTGCATCCCAATGATGAGAC 1085
 DB 206 ATGAGGAGCTGAATCTGTTATTTTGGCAAGGTTTGCATCCCAATGATGAGAC 265
 QY 1086 CTGTGAGAGAGGTGATATCCAGTTAGGTTAAGAAATTCATACAAACCTTAAGCTCAG 1145
 DB 266 CAGTGAAGGAGGATGATCCAGTTCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 325
 QY 1146 GCACCCCTTATACCAAGCAAGAGACATGATTAAGTCTGTTGTTACTACTAGCATAG 1205
 DB 326 GCACGTGATCATAAACAAGAGATATGGGCAAG-----CATATTAACCAAGCATTTG 379
 QY 1206 TGCTCAAGTCAATATGCTATGTTGGACATTTGTGACCTGGATGCTTGGTCAATG 1265
 DB 380 TCCTGAATCAATATTAATACATCTGATATATAGTGAAGCAAGATGCTGGACAGTATG 439
 QY 1266 GTTTCTGCAAGGGGATCAAGTATTTGCTATTTGAAGATCAATGATATCTGTGAT 1325
 DB 440 GCTTTCTAGCTTAAGTCTTCTCAATAT-----TTGAAGATTTGGTATCTGTGAT 493
 QY 1326 GTGTGCTACCAAGTGAAGTATGTTTCTGTGCTCACTTATCATCAAAAGATCTGAGTA 1385
 DB 494 CTGGGCTACTAGTGAAGTGAAGTCAATCATTTGACATGATCCATCAAAAGCTGGAGTTC 553
 QY 1386 GGGAACTGATACAGAGGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1445
 DB 554 GTGATATGATCCAGCAG-----GAGCTGATGATGATGATGATGATGATGATGATGAT 607
 QY 1446 CAATGTTGCTGCTACTTACCTGACAGAGGCGGATTAATTCATTAATGCAAAATGTTGA 1505
 DB 608 CAGTTGTTCACTCTCTCAAGCAGACATCAATCATCTCCCTGATGAGGAAAGCTGCAAGAT 667
 QY 1506 CGTCTGATATACATAAAGAGAGGAGCGTGTGCTGAGGAAAGTGGGGTATATGTTTACA 1565
 DB 668 CGTCTGATATTTTGAAGAGGCGTGTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 727
 QY 1566 TGATCTGCAAGAGGCGTGTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1625
 DB 728 TGATTTGCAAGGGGCGTGTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 787
 QY 1626 AGGACCTGCTGAAGAGGCGTGTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1685
 DB 788 AGCAGTGGTGAAGAGGCGTGTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 847
 QY 1686 AAGCGAGAACT 1698
 DB 848 AGGAGGCGGACCT 860

RESULT 9
 CK206100
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

CK206100 1101 bp mRNA linear EST 08-DEC-2003
 FAS017676 Triticum aestivum FGAS: Library 5 GATE 7 Triticum
 aestivum cdna, mRNA sequence.
 CK206100
 CK206100.1 GI:39568490
 EST.
 Triticum aestivum (bread wheat)
 Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooidae; Triticeae; Triticum.
 1 (bases 1 to 1101)
 Allard, F., Crosby, M.L., Danyluk, J., Eudes, F., Fick, M., Gaudet, D.,
 Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A.,
 Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, D.,
 Peniket, C., Roach, J.L. and Sarhan, F.
 Functional Genomics of Abiotic Stress in Wheat and Canola Crops
 Unpublished (2003)
 Contact: Wm L Crosby
 Bioinformatics
 University of Saskatchewan, Department of Computer Science
 1C101 Engineering Building, 57 Campus Drive, Saskatoon.

Saskatoon, S7N 5A9, Canada
 Tel: 306 966 1769
 Fax: 306 966 2033
 Email: fgas@usask.ca
 This sequence is the direct result of the Base calling software
 phred (default parameters). It is the raw base calls. To aid in the
 identification of the high quality insert the software Lucy
 (default parameters) has been run on this sequence. Lucy identified
 the region (1,853).
 Plate: L5B06 row: E column: 14.
 Location/Qualifiers
 1..1101

FEATURES
 source
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /db_xref="taxon:4565"
 /clone_lib="Triticum aestivum FGAS: Library 5 GATE 7"
 /note="vector: pCMV.SPORT6; Crown and developmental stages
 of spike formation in wheat cultivar Norstar. 4 mRNA
 populations were combined before constructing the library.
 The first mRNA population is from 1cm crown sections after
 30 days of cold acclimation. The second is from 1cm crown
 sections after 11 days of deacclimation (before
 deacclimation plants were fully vernalized for 49 days).
 The third is from different developmental stages of spike
 formation (5 to 50mm) that still have not emerged from the
 leaf (dissection required). The last is from different
 developmental stages of spike and seed formation after
 having emerged from the leaf (visible). First strand
 synthesis in this library was done in the presence of
 methylated dCTP thereby protecting from internal cleavage
 with NotI."

ORIGIN

Query Match 22.4%; Score 445.2; DB 14; Length 1101;
 Best Local Similarity 73.4%; Pred. No. 6,66-99;
 Matches 629; Conservative 0; Mismatches 209; Indels 19; Gaps 4;

QY 846 AAGGCTGGAATCTGTGCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 905
 DB 23 AAGGCTGGAATCTGTGCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 82
 QY 906 CAACTGTTGTAAGCTTGGGCTGAGAGAAATCAGTATGGAAGTGAAGTGAAGTGAAGTGAAG 965
 DB 83 CAACTGTTGTAAGCTTGGGCTGAGAGAAATCAGTATGGAAGTGAAGTGAAGTGAAGTGAAG 142
 QY 966 TACTTACTTGTATCCAAATATCTACCAATGCAAGACTGTTCCATCTTAACATTTG 1025
 DB 143 TGTGACGTGATCCAAATATTTATGCAAAATGCGGTACAGTACCTTACTGACTTTTG 202
 QY 1026 AAGAGCCACAGAACTTGTGTTATTTTGTGCTCAGTTTGCATCCCAATGATGAGAC 1085
 DB 203 ATGAGGACAGTGAATCTGTTATTTTGTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 262
 QY 1086 CTGCTAGAGAGGTGATATTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1145
 DB 263 CAGCTAGGGAAGGTGATCCAGTTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 322
 QY 1146 GCACCCCTTATACCAAGCAAGAGATGATTAAGTCTGTTTACTACTAGCATAG 1205
 DB 323 GCACGTGATCATAAACAAGAGATATGGGCAAG-----CATATTAACCAAGCATTTG 376
 QY 1206 TGCTCAAGTCAATATGCTATGTTGGACATTTGTGACCTGGATGCTTGGTCAATG 1265
 DB 377 TCCTGAATCAATATTAATACATCTGATATATAGTGAAGCAAGATGCTGGACAGTATG 436
 QY 1266 GTTTCTGCAAGGGATCAAGTATTTGCTATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1325
 DB 437 GCTTTCTAGCTTAAGTCTTCTCAATAT-----TTGAAGATTTGGTATCTGTGAT 490
 QY 1326 GTGTGCTACCAAGTGAAGTATGTTTCTGTGCTCACTTGAATCAATCAAAAGATCTGAGTA 1385
 DB 491 CTGGGCTACTAGTGAAGTGAAGTCAATCATTTGACACTGATCAATCAAAAGCTGGAGTTC 550

QY 1386 GGGAAGTGTATACGACGAGGAGTAACTTGACCATGTATGTGAAGAGCTTGAGAAAATAG 1445
 Db 551 GTGAATTGATCCGACG-----GAGCTTGATCATGTATGTGAAGAGCTTGAAAAGATTTG 604
 QY 1446 CAATTTGTGTCTACTTCAAGAGAGGCGATTAATTTCACTTATGTGAAGATGTAGCAAT 1505
 Db 605 CAGTTGTTCATCTCTCAAGACACATCATATTTCTCTGTATAGGAGAGCTGACAGAT 664
 QY 1506 CGTCTGTGATCTAGAAAAGACGAGACGTGTCTGAGAGAAAAGTGGGTTAATGTGAGA 1565
 Db 665 CGTCTGTGATCTGTGAGAGGCGTTCACGTTCATACGAGAAATGTGTATGTTCAGA 724
 QY 1566 TGATTCGAGAGAGGCGTCAAGTTAATGTCTGATAGTTCATATGTAGCAATGCA 1625
 Db 725 TGATTCGAGAGGCGTCAAGTTAATGTCTGATAGTTCATATGTAGCAATGCA 784
 QY 1626 AGGCACTCTGAGAGGCGTTCATACAGGCGTTCATTAAGAGAGATGTCTTATCAGAGTCG 1685
 Db 785 AGGAGTCCGCGAGAGCC-TCACTGCGCATTTCTTGAAGAGAGGCGCTTTGACAGAGTCG 843
 QY 1686 AAGCGAGAACCTACTC 1702
 Db 844 AGGAGCGAGACTTGAC 860

RESULT 10

CB679618

LOCUS 827 bp mRNA linear EST 09-APR-2003
 DEFINITION OSJNEF03F11.f OSJNEF Oryza sativa (japonica cultivar-group) cDNA
 clone OSJNEF03F11 5', mRNA sequence.

ACCESSION CB679618
 VERSION CB679618.1 GI:29683343

KEYWORDS

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

AUTHORS

1 (bases 1 to 827)
 Jantsuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
 Kudrna,D., Dean,R., Soderlund,C., Wang,R. and Wang,G.
 Large-scale identification of ESTs involved in the interaction
 between rice and Magnaporthe grisea

JOURNAL

COMMENT

Unpublished (2003)
 Contact: Rod Wang
 Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3667
 Fax: 520 621 9288
 Email: http://genome.arizona.edu

PCR Primers

FORWARD: gta aac cga cgg cca gtc
 BACKWARD: gga aac agc tat gac cat g
 Plate: 03 row: F column: 11
 Seq primer: gta aac cga cgg cca gtc.

FEATURES

Location/Qualifiers
 1..827

source

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="OSJNEF03F11"

/rissue_type="leaf"

/dev_stage="3 week"

/lab_host="DH10B"

/clone_lib="OSJNEF"

/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
 XhoI; Uninfected Control"

ORIGIN

Query Match

22.4%; Score 444; Db 14; Length 827;

Best Local Similarity 73.3%; Pred. No. 1,1e-98;
 Matches 610; Conservative 0; Mismatches 212; Indels 10; Gaps 3;

QY 406 CACACACCTTCCCT-TGCTGAGAGAGAGCAGTAGAGGTGTGAGTTATCATGTTCTG 464
 Db 1 CAATACCTCCTCCTCTTGGAGAGAGGTGTCACTCGGGGCCGAGAGCGCTCG 60
 QY 465 AATGAGAGTGTGATATGTCTAAAGCCTTACATCAAGAGCGGTGATGAATCTGAC 524
 Db 61 AATTCGAGGTGCACTTATCAAGAGCTTCAATGTATGACTATTTGATGACCTTGAT 120
 QY 525 TTCCAGNATCTGTATCAAGCCTT---ATGAACTGAGCACTATTAAGATATCG 581
 Db 121 TGGATAGATCATGTTCTTCAAGTTTATTTGAAGAAATGGAACAACCTTTAAGGTGTG 180
 QY 582 CTATATGAAAGCTGACGCTTGAACAGTACTTCTTTTCTATTTGGAAATGCA 641
 Db 181 CTATATGAAAGACTAACTCTGAGACAGGGAATTAACCTTGTCTTCTGATGAAATGCA 240
 QY 642 TGTCACAGAGATTTTCTGCTATTTGAACAAATTCGTGCAAGGACAGGAGATG 701
 Db 241 TGCTCAAGAAATATTTCTGATATTTGAATTAACCTTGGAAAAAGCTCGGCAATG 300
 QY 702 AGCATTGATATTTGTTTCAATTAACAATGAAATTTGTAATGCGGATATCTTAAG 761
 Db 301 ATGATTTGATATTTGCTTTTATTAATCTAGTATTTCAAAATGCGGACATTTTGAAG 360
 QY 762 CAACCTATCCGCTGTTGCGAGAGACTTATGAGGCGCTGATACAGAGATCCAGCATAC 821
 Db 361 CCAATATCTGCTGTTGCAAGAGGTGATGAGGTGATGATGACAGCCCTGCTATTC 420
 QY 822 CTGTTGTACTGAGTTCCTTGGAGAGGCTGGAATGTGCTGTGTAATCTTAAAGCC 881
 Db 421 CTATATGATACGTTCTTCTTGAAGAGATGGAATCATGTGCTGACACACTTTAGGA 480
 QY 882 GAGGTGTAGTACTTACTGCTTCAACCATTTGTTAAAGCTTGGAGACGAGAAATTC 941
 Db 481 GAGGTGTAGTACTTACTGCTTCAACCATTTGTTAAAGCTTGGAGACGAGAAATTC 540
 QY 942 AGGTATGAAAGATGTTATGTTGTTACTTCTGATCCAAATATCTACCCACATGCAA 1001
 Db 541 AGGTATGAAAGATGTTATGTTGTTACTTCTGATCCAAATATCTACCCACATGCAA 600
 QY 1002 AGCTGTTCACTTAAATTTGAAGAGCCACAGACTTCTTATTTGTGCTCAGG 1061
 Db 601 TACCTGTACCTTACTGCTTTTGAAGAGGAGTGAACCTGCTTCTTGTGACAGG 660
 QY 1062 TTTGATTCACAAATGATGAGAGCTGCTGAGAGAGGTGATTTCCAGTTAGGTTAAG 1121
 Db 661 TTTTGCATCCCATCATGATGACACAGCAAGAGGTGATATACAGAGTTAGGTTAAG 720
 QY 1122 ATTATACACCCCTTAAGCTCAGGACCTTATTTACAGACAAAGACATGATGATG 1181
 Db 721 ATTATATACCCGTCAGACACCTGTTATTTATTAACAAAGAGATTTAGCAAG 780
 QY 1182 GTCTGTGTACTACTGATAGTGTGCTCAAGTCAATGTGATGATTTGGA 1233
 Db 781 G-----TATATTAACGACATTTATTAAGTCAACATTTACTATGCTGGA 826

RESULT 11

BO506340

LOCUS 792 bp mRNA linear EST 07-MAR-2003
 DEFINITION EST131755 Generation of a set of potato cDNA clones for microarray
 analyses mixed potato tissues Solanum tuberosum cDNA clone STM0371

ACCESSION

BO506340

VERSION

BO506340.2

KEYWORDS

SOURCE

ORGANISM

Solanum tuberosum (potato)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE
AUTHORS

asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 792)
Buell, C.R., Hart, A., Baker, B., Tanksey, S., Fry, W., Smart, C.,
Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and
Karamycheva, S.A.
Generation of a set of potato cDNA clones for microarray analyses
Unpublished (2002)
On Jun 10, 2002 this sequence version replaced gi:21365209.
Contact: Robin Buell
The Institute for Genomic Research
3712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igf.rockefeller.edu
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T3.

FEATURES
source

Location/Qualifiers
1..792
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec or Binje"
/db_xref="taxon:4113"
/clone="SIMGJ71"
/tissue_type="mixed tissues"
/lab_host="SOLAR"
/clone_lib="Generation of a set of potato cDNA clones for
microarray analyses mixed potato tissues"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Combination of untreated and phytophthora
infestans-treated libraries of stolons, leaves, leaflets,
axillary buds of stem explants, petioles, germinating
eyes, tubers, or roots."

ORIGIN

Query Match 22.1%; Score 436.6; DB 13; Length 792;
Best Local Similarity 76.1%; Pred. No. 2.4e-97;
Matches 556; Conservative 0; Mismatches 169; Indels 6; Gaps 1;
QY 554 GAACCTGAGCACTATTTGAAGATTCGCTATGATGAAGAAGCTAGCCCTAGAGCAAGT 613
DB 63 GAACCTGAGCACTATTTGAAGATTCGCTATGATGAAGAAGCTAGCCCTAGAGCAAGT 122
QY 614 GACTACCTGTTTCATTTGAGAGATGATGTCACACAGATTTTTCCTGATTTTGAAC 673
DB 123 GACTACCTGTTTCATTTGAGAGATGATGTCACACAGATTTTTCCTGATTTTGAAC 182
QY 674 AAAATTCGTGCAAGGACGGCAGTATGACGATTTGATTTGTTTCAATTCACATGAT 733
DB 183 AAAATTCGTGCAAGGACGGCAGTATGACGATTTGATTTGTTTCAATTCACATGAT 242
QY 734 GAATTCGATATCGGATATCTTGAAGCAACTTATCCGCTGTTGCAAGAGACTTCAAT 793
DB 243 GAATTCGATATCGGATATCTTGAAGCAACTTATCCGCTGTTGCAAGAGACTTCAAT 302
QY 794 GGGGACTGATACAGATCCAGGATACCTGTTTACTGCGGTTCTTGGGAAGGCTGG 853
DB 303 GGGGACTGATACAGATCCAGGATACCTGTTTACTGCGGTTCTTGGGAAGGCTGG 362
QY 854 AAATCTGGGCTGTAATCTTGAAGCCAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 913
DB 363 AGAAGCTTGTCTGTGATACCTAGTGAAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 422
QY 914 GGTAAAGCTTGGAGTGAAGAAATTCAGGTATGAAAGATGTTGATGTTGATTTACT 973
DB 423 GGTAAAGCTTGGAGTGAAGAAATTCAGGTATGAAAGATGTTGATGTTGATTTGACA 482
QY 974 TGGATGCAATATCTTACCAATGCAAGAGAGTGTCCATCTTAACTTAACTTAAAGGCC 1033
DB 483 TGGATGCAATATCTTACCAATGCAAGAGAGTGTCCATCTTAACTTAACTTAAAGGCC 542
QY 1034 ACAAGACTTCTTATTTGTGCTCAGGTTTGGATTCACAAATGATGAGAGCTTCTAGA 1093
DB 543 GCTAACTGCAATCTTGGGCGCAGAGTCTTCATCCACAGATCCATGAGGCGCTCTAGA 602

QY 1094 GAAGGTGATATTCGATTTAGGTTTGAAGATTCATACCACTTAAGCTCCAGGCACTT 1153
DB 603 GAGGCTGATATTCCTGTTAGGTTTAAAGATTCGATTAATCCTTAAGCTCCGCACTTC 662
QY 1154 ATTACCAACAAAGAGATGATGAATTAAGCTGTTGTTTACTACTGATGATGCTCAAG 1213
DB 663 ATCGTTTAAACAGAGATGATGAATTAAGCTGTTGTTTACTACTGATGATGCTCAAG 716
QY 1214 TCAATGCTACTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1273
DB 717 AGCATGATACCATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 776
QY 1274 GCAAGGATATC 1284
DB 777 CAAGGATATC 787

RESULT 12

CF449571 764 bp mRNA linear EST 04-SEP-2003
LOCUS EST85916 normalized cDNA library of onion Allium cepa cDNA clone
DEFINITION ACACU05, mRNA sequence.
ACCESSION CF449571
VERSION CF449571.1 GI:34472273
KEYWORDS EST
SOURCE Allium cepa (onion)
ORGANISM Allium cepa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
Allium.

REFERENCE 1 (bases 1 to 764)
AUTHORS Havey, M.J., Chung, F., Van Aken, S., Uteback, T. and Town, C.D.
TITLE Expressed Sequence Tags from a normalized library of mixed onion
tissues (Allium cepa)
JOURNAL Unpublished (2003)
COMMENT Contact: Havey MJ
Department of Horticulture
USDA-ARS and University of Wisconsin
1575 Linden Drive, Madison, WI 53706, USA
Tel: 608-262-1830
Fax: 608-262-4743
Email: mjhvey@facstaff.wisc.edu
TIGR sequence name ACACU05TR. For more information:
http://haveylab.hort.wisc.edu
Seq primer: CAG GAA ACA GCT ATG ACC.

FEATURES

source
1..764
Location/Qualifiers
/organism="Allium cepa"
/mol_type="mRNA"
/cultivar="Red Creole (bulbs), unknown (callus), Ebano &
Texas Legend (roots)"
/db_xref="taxon:4679"
/clone="ACACU05"
/tissue_type="Callus, roots, and young bulbs"
/clone_lib="normalized cDNA library of onion"
/note="Vector: pCMVSPORT6.1-cdbd (Invitrogen); Site_1:
EcoRV (5'); Site_2: NotI (3'); Equal molar amounts of mRNA
from callus, roots, and young bulbs were combined to
synthesize the library. Normalization to enrich for
low-copy transcripts was performed by proprietary
techniques of Invitrogen."

ORIGIN

Query Match 22.0%; Score 437.6; DB 14; Length 764;
Best Local Similarity 74.8%; Pred. No. 4.2e-97;
Matches 579; Conservative 0; Mismatches 184; Indels 11; Gaps 2;
QY 581 GCTATGATGAAGAGCTTACCTGATGATGATGATGATGATGATGATGATGATGATGAT 640
DB 1 GCTATGATGAAGAGCTTACCTGATGATGATGATGATGATGATGATGATGATGATGAT 60
QY 641 ATGTCACCAAGAGATTTTTCCTGATTTTGAACAAATTCGTGCAAGGACGCACTAT 700

Db	Accession	Definition	LOCUS	Accession	Version	Keywords	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	61	ATGCTACAGAGATTTTGGCGCTTACTGAAACAAATTTGTCAGAAAGCAGGAGAT	120										
Qy	701	GACGCAATTGATTTGGTTTCATTACACATGATGATGATTTGGTATGCGGATTCCTTGA	760										
Db	121	GATGCTTCGATATTTGGTATCATACACAGATGATTTTACAAATGCTGACATCTTGAA	180										
Qy	761	GCAACCTATCCTGCTGTTGSCAAGAGACTTCATGSGGAGCTGATACAGATCCAGAT	820										
Db	181	GCACTTATCCTTCTGTGCAAAAAGATGATGATGAGATTGGATGCTAATCTTCAT	240										
Qy	821	CCTGTTGTTACTGCGGCTTCTTGGAAAGGCTGGAATCTGTGCTGTAACTACTT	880										
Db	241	CCAGTTGTAAACGGGCTTTTGGAAAGGCTGGAATCTGTGCTGTAACTACTT	300										
Qy	881	CGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	940										
Db	301	AGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	360										
Qy	941	CAGGTATGGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT	1000										
Db	361	CAGGTATGGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT	420										
Qy	1001	AAGACTGTCCACTACTTAACTTGAAGAGCCACAGAACTTGTATTTGTTG	1060										
Db	421	AGACCTGTCCACTACTTAACTTGAAGAGCCACAGAACTTGTATTTGTTG	480										
Qy	1061	GTTTTCATCCCATCATCATCATCATCATCATCATCATCATCATCATCATCAT	1120										
Db	481	GTTTTCATCCCATCATCATCATCATCATCATCATCATCATCATCATCATCAT	540										
Qy	1121	AATTCATCAACCCCTTAACTTCAAGCACCCTTATTCACAGAAAGACATGAT	1180										
Db	541	AATTCATCAACCCCTTAACTTCAAGCACCCTTATTCACAGAAAGACATGAT	594										
Qy	1181	GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1240										
Db	595	ACCGGGGCTGATTAACACAGATGCTTGAATCAATGATTAACATGCTGAT	654										
Qy	1241	AGCAGCTCGATCTGTCAGATGCTTTCGCAAGGATACAGATTTGCTAT	1300										
Db	655	AGCAGCTCGATCTGTCAGATGCTTTCGCAAGGATACAGATTTGCTAT	714										
Qy	1301	GAAAGCTATGATATCTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1354										
Db	715	AATGATAT-----TTTCGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG	763										
RESULT 13	BG053377	548 bp	MRNA	linear	EST 25-JAN-2001								
LOCUS	BG053377	RHIZ2_26_H10_b1_A003	Rhizome2 (RHIZ2)	sorghum propinquum	cdNA, mRNA								
DEFINITION	sequence												
ACCESSION	BG053377												
VERSION	BG053377.1	GI:12508988											
KEYWORDS	EST.												
SOURCE	Sorghum propinquum												
ORGANISM	Sorghum propinquum												
REFERENCE	1	(bases 1 to 548)											
AUTHORS	Cordomier-Pratt, M.-M., Gingle, A., Paterson, A., Sudman, M. and Pratt, L.H.												
TITLE	An EST database from Sorghum: Sorghum propinquum rhizomes												
JOURNAL	Unpublished (2000)												
COMMENT	Contact: Cordomier-Pratt MM Laboratory for Genomics and Bioinformatics The University of Georgia, Department of Plant Biology Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA Tel: 706 542 1860 Fax: 706 583 0210 Email: mmpratt@uga.edu	</											

FEATURES					
	source	location/qualifiers			
		1..548			
		/organism="Scorpius propinquum"			
		/mol_type="mRNA"			
		/db_xref="taxon:132711"			
		/clone_lib="Rhizome2 (RH122)"			
		/note="Organ: Rhizomes; Vector: pBluescript II from Lambda Zap II; Site 1: XhoI; Site 2: EcoRI. The library was made from poly-A RNA in the cloning vector lambda Zap II.			
		Clones to be sequenced were prepared by mass excision."			
 ORIGIN					
	Query Match	22.0%; Score 436.2; DB 10; Length 548;			
	Best Local Similarity	90.6%; Pred. No. 7.9e-97;			
	Matches 465; Conservative 0; Mismatched 48; Indels 0; Gaps 0;				
Qy	4	CCAGCAGAGGAGGAGGACTTCAAAATGGCAATCCAGATGGCGCGGCCGCCGC	63		
Dd	36	CCCCACGAGTAGAGGGGCCAAGAAATGGCGACCACAATGCCATGGCGCCGACACG	95		
Qy	64	CCGCGCTCGTCCTCGTGATACCTCCGCGAGCTTTGACATGTTCAGAGACTGGCGTCTT	123		
Dd	96	CCGCGCTCGTCCTCGTGATACCTCCGCGAGCTTCGGACATGTTCACAACTGACGCTGCTT	155		
Qy	124	CGGTACCCGGAACCGGCGCTCGCGGTGTAAGAAGGTTGTCAATGGTGGTCCGCACTCCAC	183		
Dd	156	CGATTCCGGAACCGAGCTTTCGATGCCAGAGGTTGTCAATGGTGGTCCGCACTCCAC	215		
Qy	184	CAGCCGTCCGAGCCCAAAGCAGCGAGCGGAGAGCGGCGTCTTGGGCGCGCTGTTCTCG	243		
Dd	216	CCGCGCGTCCGAGCCCAAAGTATAGGAGAGCGAGCGGAGAGCGGCGTCTTGGTCTCGG	275		
Qy	244	AGGGCTCCGGAATGAGAGGGATTTGGGGGATTCAGCTCACGCGTGTGATGAAGTTCCGGAGGCTC	303		
Dd	276	AGGGCTCCGGAATGAGGAGATTTGGGGGATTCAGCTCACGCGTGTGATGAAGTTCCGGAGGCTC	335		
Qy	304	CTCGGTGTCGTGCGCCGCGAGAGATGCGTAGAGTGCGCGGCGCTTCATCTTGAAGCTTCCCGA	363		
Dd	336	CTCAGTGGCGTCGCGCCGCGAGATGCGTAGAGTGCGCGGCGCTTCATCTTGAAGCTTCCCGA	395		
Qy	364	GGAGGCGCCCGCTCGCTGCTCTCTGTCGCAAGGGGAAAAACACCAACAACTTCTCTCTTGC	423		
Dd	396	GAGAGCGCCCGCTCGCTGCTCTCTGTCGCAAGGGGAAAAACACCAACAACTTCTCTCTTGC	455		
Qy	424	TGAGAGGAAGGAGTAGGGGTGTGAGATTATCCATGTTTTCTGAATCGAAAGTGGATAT	483		
Dd	456	TGAGAGGAATGAGTAGGGGTGTGAGATTATCCATGTTTTCTGAATCGAAAGTGGATAT	515		
Qy	484	GGTCAAAAGCCATCATCAAGACGGTGAATGA	516		
Dd	516	GATCAAAAGCTCATCATCAAGACGAGGATGA	548		
 RESULT 14					
BQ969663		727 bp mRNA linear EST 21-AUG-2002			
LOCUS		clone OHB39D05, mRNA sequence.			
DEFINITION		OHB39D05.yv.ab1 CH ABCDI sunflower RH4801 Helianthus annuus cDNA			
ACCESSION		BQ969663.1 GI:22387184			
VERSION		EST.			
KEYWORDS		Helianthus annuus (common sunflower)			
SOURCE		Helianthus annuus			
ORGANISM		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae; Helianthus.			
REFERENCE		1 (bases 1 to 727) Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,			

Elison, P., Kolman, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,
Lai, Z., Church, S., Jackson, L., and Bradford, K.
Lecture and Sunflower ESTs from the Composite Genome Project
http://composita.ucdavis.edu/
Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmor]
Department of Vegetable Crops, R.W.Michelmor Lab
University of California at Davis (UCD)
Ammundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]
singleton, see http://cgpdb.ucdavis.edu/ for details.
Plate: QHB39 row: D column: 05.

FEATURES

Source

1..727
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA801"
/db_xref="taxon:4232"
/clone="QHB39D05"
/lab_host="E.coli"
/clone_lib="QH ABCDI sunflower RHA801"
/note="Vector: pBRCDNA51AB; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_LIB-QH ABCDI sunflower RHA801
TAG_SEQ=TCGCACGCG"

ORIGIN

Query Match 21.9%; Score 434.2; DB 13; Length 727;

Best Local Similarity 75.6%; Pred. No. 2.9e-96;

Matches 554; Conservative 0; Mismatches 173; Indels 6; Gaps 1;

578 ATGCGATGATGAAAGAGCGAGCGCTAGACAGTACCTGTTTCATTGGAGAA 637
1 ATTCTGTACTTAAGAAATTCACACCGCGAAGAGATTACTTACTCTGTTGGAG 60
638 TGCATGTCACACGAGATTTTCTGCTTAATTGAACAAATTGCTGTCAGGACGCGAG 697
61 TGCATGTCACACGAGATTTTCTGCTTAATTGAATTAATCGGTGTCAGGACGCGAA 120
698 TATGACGCACTTATGTTGGTTTCATTACACGATGATTTGGTAATCGGATCTTA 757
121 TATGATGCAATTCGAAATTTGTTTCATACACGATGATTTTCTTAATCGGATTTTGG 180
758 GAAGCAACCTATCTGCTGTTGGAGAGACTTCATGAGGATGATACAGATCCAGCG 817
181 GAAGCAACCTATCTGCTGTTGGAGAGACTTCATGAGGATGATACAGATCCAGCG 240
818 ATACCTGTTGTTACTGGTTCCTTGGAGAGGCTGGAATTCGTGCTCTACTACTTTA 877
241 ATTCTATTTGTCAGTCTGCTTCTTGAAGAGGTTGAGAACTTGCTGCTGATCAGC 300
878 GCGCAGAGTGTAGTACTGCTGCTACACCATTTGTAAGCTTGGAGTGAAGAA 937
301 GCGCAGAGTGTAGTACTGCTGCTACACCATTTGTAAGCTTGGAGTGAAGAA 360
938 ATTGAGGTATGAAAGATTTGATGCTGCTACTTGTGATCCAAATTCATACCAAT 997
361 ATTGAGGTATGAAAGATTTGATGCTGCTACTTGTGATCCAAATTCATACCAAT 420
998 GAAAGACTGTTCCATCTTAATCAATTTGAAGAGCCACAGAACTTGCTTAATTTGGTGT 1057
421 GCTGACCTGCTCCATTTTGAAGCTTGAAGAGCCACAGAACTTGCTTAATTTGGTGT 480

QY 1058 CAGTTTTCATCCACATTCAGTATGAGACCTCTAGAGAGGTATATCCAGTTAGGTT 1117
DB 481 CAGTCTTACATCCACATTCAGTATGAGACCTCTAGAGAGGTATATCCAGTTAGGTT 540
QY 1118 AAGATTCATACACCTTAAGCTTCAGGACCTTATTTACGACAAAGAGATGAT 1177
DB 541 AAAAAATGTTATACCCCAATGCTCTGTTACCTTAATACCAATCAAGAGATGAT 600
QY 1178 AANGCTGCTGTTACTACTAGCATAGTCTCAAGTCAATGCTACTATGTTGACAT 1237
DB 601 AA-----GCGATCTACACGATGTTTGAAGTATGTAACCATGTTAGACAT 654
QY 1238 GTGACACTCGATCTGTTGATGATGTTTCTGCAAGGATATCAGATTTGCTAT 1297
DB 655 GTTACGACTCGATCTGTTGATGATGTTTCTGCAAGGATTTTGAATTTGAG 714
QY 1298 ATTGAAGATCTAT 1310
DB 715 GATTGGGTTAT 727

RESULT 15
BU988067 607 bp mRNA linear EST 22-OCT-2002
LOCUS HPI6004 HF Hordeum vulgare subsp. vulgare cDNA clone HPI6004
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION BU988067
VERSION BU988067.1 GI:24239013
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Striptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 607)
Radtchuk, V., Zhang, H., Weschke, W., Potokina, E. and Webus, U.
Barley ESTs from developing seeds
Unpublished (2002)
JOURNAL Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 607 Std Error: 0.00
Plate: 16 Row: J Column: 4
Seq primer: MJREV.

FEATURES

Source

1..607
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"
/db_xref="GABI:248011"
/db_xref="taxon:112509"
/clone="HPI6J04"
/tissue_type="caryopsis"
/dev_stage="developing caryopsis, 16-25 DAF (days after
flowering)"
/lab_host="XLI0-GOLD"
/clone_lib="HF"
/note="Vector: pBluescript SK+, Site 1: EcoRI (5'-end of
cDNA); Site 2: XhoI (3'-end of cDNA); developing
caryopsis, 16-25 DAF(days after flowering) Due to a cloning
artefact caused by the kit, in most cases the EcoRI site
is NOT present, as well as the EcoRI adapter used for
cloning. To excise the insert, restriction sites upstream
EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also
due to the cloning system used blue/white selection for
recombinants is not 100% reliable. Average insert size is
940 bp"

ORIGIN

Query Match 21.7%; Score 431; DB 13; Length 607;
Best Local Similarity 84.3%; Pred. No. 1,6e-95;
Matches 527; Conservative 0; Mismatches 80; Indels 18; Gaps 3;

```
QY 893 GACTTGACCTGCTACCAACCTTGAAGCTTGGAGCTGAGAGAAATTCAGGTATGGAA 952
   |||||
Db 1 GACTTGACCTGCTACCAACCTTGAAGCTTGGAGCTGAGAGAAATTCAGGTATGGAA 60
   |||||

QY 953 GATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1012
   |||||
Db 61 GATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
   |||||

QY 1013 TACTTACATTTTGAAGGCGCACAGACTGTGTTTGGTGGCTCAGGTTTGATCCA 1072
   |||||
Db 121 TACTTACATTTTGAAGGCGCTGCTGATCTTATTTTGGTGGCTCAGGTTTGATCCA 180
   |||||

QY 1073 CAATCGATGAGACCTGCTAGAGAGGTGATTCAGTTAGGGTTAAGAAATTCATCAAC 1132
   |||||
Db 181 CAGTCATGCGACCTGCTAGAGAGGTGATTCAGTTAGGGTTAAGAAATTCATCAAC 240
   |||||

QY 1133 CCTAAGCTCCAGGCGACCTTATTAACAACAACAAGACATGGAATTAAGGCTGTTGTA 1192
   |||||
Db 241 CCTAAGCTCCAGGCGACCTTATTAACAACAACAAGACATGGAATTAAGGCTGTTGTA 294
   |||||

QY 1193 CTACTAGCATAGTGTCTCAAGTCAATGCACTATGTTGACATTTGTGAGCACTGGATG 1252
   |||||
Db 295 CTACTAGCATAGTGTCTCAAGTCAATGCACTATGTTGACATTTGTGAGCACTGGATG 354
   |||||

QY 1253 CTTGGTCAGTATGTTTCTGGCAAGGGTATCAGGATTTGCTATATTGAAGATCTATGT 1312
   |||||
Db 355 CTTGGTCAGTATGTTTCTGGCAAGGGTATCAGGATTTGCTATATTGAAGATCTATGT 408
   |||||

QY 1313 AATCTGTGATTTGTTGTTCTACAGTGAAGTTAGTTTCTGTCTCACTGATCCATCA 1372
   |||||
Db 409 AATCTGTGATTTGTTGTTCTACAGTGAAGTTAGTTTCTGTCTCACTGATCCATCA 468
   |||||

QY 1373 AAGATCTGAGTAGGGAATCTATACAGCAGGCAAGTGAATTGACCATGTTGAAGAG 1432
   |||||
Db 469 AAGATCTGAGTAGGGAATCTATACAGCAGGCAAGTGAATTGACCATGTTGAAGAG 522
   |||||

QY 1433 CTTGAGAAATATAGCAATTTGTTGCTACTTACAGAGGCGCATATTTCACTTATCGGA 1492
   |||||
Db 523 CTTGAGAAATATAGCAATTTGTTGCTACTTACAGAGGCGCATATTTCACTTATCGGA 582
   |||||

QY 1493 AATGTGAGCAATGCTCTGATAC 1517
   |||||
Db 583 AATGTGAGCAATGCTCTGATAC 607
   |||||
```

Search completed: March 23, 2004, 11:04:30
Job time : 5272 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 11:04:35 ; Search time 77 Seconds
(without alignments)
2054.891 Million cell updates/sec

Title: US-09-890-813-6
Perfect score: 2813
Sequence: 1 MAIPVRSAAAPRLVPSIFP.....QAFEDDVLSQVEARNLVG 560

Scoring table: BLASTUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum March 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	% Length	ID	Description
1	2809	99.9	560	AAE04353	AAE04353 Zea mays
2	2678.5	95.2	555	AAE04354	AAE04354 Zea mays
3	1835	65.2	564	AAE04359	AAE04359 Glycine m
4	1789	63.6	569	ABE93449	ABE93449 Herbicida
5	1747.5	62.1	544	AAE04236	AAE04236 Arabidops
6	1747.5	62.1	544	ABE93460	ABE93460 Herbicida
7	1728	61.4	563	AAE04215	AAE04215 Arabidops
8	1727.5	61.4	459	AAE04237	AAE04237 Arabidops
9	1708	60.7	478	AAE04216	AAE04216 Arabidops
10	1675.5	59.6	446	AAE04238	AAE04238 Arabidops
11	1675.5	59.6	555	AAE05075	AAE05075 Arabidops
12	1675.5	59.6	559	AAE05074	AAE05074 Arabidops
13	1675.5	59.6	559	ABE92174	ABE92174 Herbicida
14	1666.5	59.2	473	AAE05076	AAE05076 Arabidops
15	1656	58.9	465	AAE046217	AAE046217 Arabidops
16	1629.5	57.9	439	AAE04358	AAE04358 Wheat wr1
17	1629.5	57.9	439	AAE04358	AAE04358 Wheat wr1
18	1074.5	39.3	307	AAE04354	AAE04354 Arabidops
19	1074.5	38.2	354	AAE04330	AAE04330 Arabidops
20	1054.5	37.5	269	AAE04332	AAE04332 Arabidops
21	890.5	31.7	281	AAE04356	AAE04356 Oryza sat
22	676	24.0	215	AAE04352	AAE04352 Arabidops
23	666	23.7	152	AAE04352	AAE04352 Arabidops
24	592	21.0	465	AAE04352	AAE04352 Arabidops
25	572	20.3	449	AAE04352	AAE04352 Arabidops

26	572	20.3	449	2	AAE04353	AAE04353 Zea mays
27	572	20.3	449	2	AAE04353	AAE04353 Zea mays
28	572	20.3	449	2	AAE04353	AAE04353 Zea mays
29	572	20.3	449	2	AAE04353	AAE04353 Zea mays
30	572	20.3	449	2	AAE04353	AAE04353 Zea mays
31	572	20.3	449	2	AAE04353	AAE04353 Zea mays
32	572	20.3	449	2	AAE04353	AAE04353 Zea mays
33	572	20.3	449	2	AAE04353	AAE04353 Zea mays
34	572	20.3	449	2	AAE04353	AAE04353 Zea mays
35	572	20.3	449	2	AAE04353	AAE04353 Zea mays
36	572	20.3	449	2	AAE04353	AAE04353 Zea mays
37	572	20.3	449	2	AAE04353	AAE04353 Zea mays
38	572	20.3	449	2	AAE04353	AAE04353 Zea mays
39	572	20.3	449	2	AAE04353	AAE04353 Zea mays
40	572	20.3	449	2	AAE04353	AAE04353 Zea mays
41	572	20.3	449	2	AAE04353	AAE04353 Zea mays
42	572	20.3	449	2	AAE04353	AAE04353 Zea mays
43	572	20.3	449	2	AAE04353	AAE04353 Zea mays
44	572	20.3	449	2	AAE04353	AAE04353 Zea mays
45	572	20.3	449	2	AAE04353	AAE04353 Zea mays

ALIGNMENTS

RESULT 1
ID AAE04353 standard; protein; 560 AA.

AAE04353;

04-SEP-2001 (first entry)

Zea mays choic.pK002.k6 (FIS) clone aspartate kinase.

Corn; aspartate kinase; cell transformation; transgenic plant;
choic.pK002.k6 clone.

Zea mays.

Key Location/Qualifiers

Misc-difference 168 /label= Unknown

Misc-difference 384 /note= "Encoded by AGN"

Misc-difference 384 /note= "Encoded by AAN"

MO200146393-A2.

28-JUN-2001.

19-DEC-2000; 2000WC-US034396.

21-DEC-1999; 99US-0172944P.

(DUPO) DU POINT DE NEMOURS & CO E. I.

Falco SC, Famodu OO, Thorpe CJ;

WPI; 2001-418057/44.

N-PSDB; AAD08621.

Novel nucleic acid molecules encoding aspartate kinase useful in cell transformation and transgenic plant production procedures.

Claim 21; Fig 1; 64pp; English.

The present invention relates to nucleic acid molecules encoding aspartate kinase useful in cell transformation and transgenic plant production procedures. Nucleic acid encoding aspartate kinase is useful for producing a transgenic plant, by transforming a plant cell with a chimeric gene comprising nucleic acid encoding aspartate kinase operably

CC linked to a regulatory sequence and regenerating a plant from the
 CC transformed plant cell. The present sequence is cdpic.pk002.k6 (FIS)
 CC clone Zea mays aspartate kinase
 XX

Sequence 560 AA;
 Query Match 99.9%; Score 2809; DB 4; Length 560;
 Best Local Similarity 100.0%; Pred. No. 1,36-256;
 Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAIPVSAAPRRLLVPSIPASSGVRGLACFGRTPGRARGLSMVVADSTRRAKQAD 60
 DB 1 MAIPVSAAPRRLLVPSIPASSGVRGLACFGRTPGRARGLSMVVADSTRRAKQAD 60
 QY 61 GGDGVGAPVGLGMEGLGQDLSVYMKFGSSSVSAAARAEVAGLITFPEERPVVLS 120
 DB 61 GGDGVGAPVGLGMEGLGQDLSVYMKFGSSSVSAAARAEVAGLITFPEERPVVLS 120
 QY 121 AMGKTNNLLAGEKAVCGVIHVSEIEENMVKSLLIKTVDELGLPRSVIQDLDELQ 180
 DB 121 AMGKTNNLLAGEKAVCGVIHVSEIEENMVKSLLIKTVDELGLPRSVIQDLDELQ 180
 QY 181 LKGIAMMKELTPRTSDVLSFGECMSSTRIFSAYLKIRKAROYDAFDIGFITTDEFGNA 240
 DB 181 LKGIAMMKELTPRTSDVLSFGECMSSTRIFSAYLKIRKAROYDAFDIGFITTDEFGNA 240
 QY 241 DILEATYPAVAKRLHGDWIODPAIPVYTGFLGKWKSGAVTTLGRGSDLTATTIGKALG 300
 DB 241 DILEATYPAVAKRLHGDWIODPAIPVYTGFLGKWKSGAVTTLGRGSDLTATTIGKALG 300
 QY 301 LREIOWKVDVGLTCDPNITYPHAKTVPYLTFEAEATELAFGAQVLAHQSMRPARBGDI 360
 DB 301 LREIOWKVDVGLTCDPNITYPHAKTVPYLTFEAEATELAFGAQVLAHQSMRPARBGDI 360
 QY 361 VRVNSYNPKAPGTLITRRDMDXGLVLTSLVLSKSNVTMLDIVSTRMLGOYGFILARVSG 420
 DB 361 VRVNSYNPKAPGTLITRRDMDXGLVLTSLVLSKSNVTMLDIVSTRMLGOYGFILARVSG 420
 QY 421 ICYIEDLCISVDCVATSEVSVSLDPSKISRELIQASSELDHVVELEKIAIVRLIQ 480
 DB 421 ICYIEDLCISVDCVATSEVSVSLDPSKISRELIQASSELDHVVELEKIAIVRLIQ 480
 QY 481 RAIIISLIGNVEQSSLIIEKTGRVLRKSGVNVQVMSIQASAKVMSLIIVHSDAKALVEALH 540
 DB 481 RAIIISLIGNVEQSSLIIEKTGRVLRKSGVNVQVMSIQASAKVMSLIIVHSDAKALVEALH 540
 QY 541 QAFEDDVLSCVEAENLLVG 560
 DB 541 QAFEDDVLSCVEAENLLVG 560

RESULT 2

AAE04354 standard; protein; 555 AA.

XX AAE04354;
 AC 04-SEP-2001 (first entry)
 DT Zea mays cdpic.pk010.k1 (FIS) clone aspartate kinase.
 XX Zea mays.
 OS Zea mays.
 XX MO200146393-A2.
 XX 28-JUN-2001.
 XX 19-DEC-2000; 2000MO-US034396.
 XX 21-DEC-1999; 99US-0172944P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.
 PA Falco SC, Farnou OO, Thorpe CJ;
 PI WPI; 2001-418057/44.
 DR N-PSDB; AAD08622.
 XX Novel nucleic acid molecules encoding aspartate kinase useful in cell
 PT transformation and transgenic plant production procedures.
 CC Claim 21; Fig 1; 64pp; English.

CC The present invention relates to nucleic acid molecules encoding
 CC aspartate kinase useful in cell transformation and transgenic plant
 CC production procedures. Nucleic acid encoding aspartate kinase is useful
 CC for producing a transgenic plant, by transforming a plant cell with a
 CC chimeric gene comprising nucleic acid encoding aspartate kinase operably
 CC linked to a regulatory sequence and regenerating a plant from the
 CC transformed plant cell. The present sequence is cdpic.pk010.k1 (FIS)
 CC clone Zea mays aspartate kinase
 XX

Sequence 555 AA;

Query Match 95.2%; Score 2678.5; DB 4; Length 555;
 Best Local Similarity 96.8%; Pred. No. 36-244;
 Matches 543; Conservative 1; Mismatches 10; Indels 7; Gaps 4;

QY 1 MAIPVSAAPRRLLVPSIPASSGVRGLACFGRTPGRARGLSMVVADSTRRAKQAD 60
 DB 1 MAIPVSAAPRRLLVPSIPASSGVRGLACFGRTPGRARGLSMVVADSTRRAKQAD 60
 QY 61 GGDGVGAPVGLGMEGLGQDLSVYMKFGSSSVSAAARAEVAGLITFPEERPVVLS 120
 DB 61 GGDGVGAPVGLGMEGLGQDLSVYMKFGSSSVSAAARAEVAGLITFPEERPVVLS 120
 QY 121 AMGKTNNLLAGEKAVCGVIHVSEIEENMVKSLLIKTVDELGLPRSVIQDLDELQ 180
 DB 121 AMGKTNNLLAGEKAVCGVIHVSEIEENMVKSLLIKTVDELGLPRSVIQDLDELQ 180
 QY 180 LKGIAMMKELTPRTSDVLSFGECMSSTRIFSAYLKIRKAROYDAFDIGFITTDEFGN 239
 DB 180 LKGIAMMKELTPRTSDVLSFGECMSSTRIFSAYLKIRKAROYDAFDIGFITTDEFGN 240
 QY 240 ADILEATYPAVAKRLHGDWIODPAIPVYTGFLGKWKSGAVTTLGRGSDLTATTIGKAL 300
 DB 240 ADILEATYPAVAKRLHGDWIODPAIPVYTGFLGKWKSGAVTTLGRGSDLTATTIGKAL 300
 QY 300 GIREIOWKVDVGLTCDPNITYPHAKTVPYLTFEAEATELAFGAQVLAHQSMRPARBGDI 359
 DB 300 GIREIOWKVDVGLTCDPNITYPHAKTVPYLTFEAEATELAFGAQVLAHQSMRPARBGDI 360
 QY 360 PVVNSYNPKAPGTLITRRDMDXGLVLTSLVLSKSNVTMLDIVSTRMLGOYGFILARVSG 419
 DB 360 PVVNSYNPKAPGTLITRRDMDXGLVLTSLVLSKSNVTMLDIVSTRMLGOYGFILARVSG 418
 QY 419 AII--FEDLCISVDCVATSEVSVSLDPSKISRELIQ--ELDHVVELEKIAIVRLIQ 474
 DB 419 AII--FEDLCISVDCVATSEVSVSLDPSKISRELIQ--ELDHVVELEKIAIVRLIQ 474
 QY 480 ORAIISLIGNVEQSSLIIEKTGRVLRKSGVNVQVMSIQASAKVMSLIIVHSDAKALVEAL 539
 DB 475 ORAIISLIGNVEQSSLIIEKTGRVLRKSGVNVQVMSIQASAKVMSLIIVHSDAKALVEAL 534
 QY 540 HQAFEDDVLSCVEAENLLVG 560
 DB 535 HQAFEDDVLSCVEAENLLVG 555

RESULT 3

AAE04359 standard; protein; 564 AA.

XX

AAE04359;
 04-SEP-2001 (first entry)
 Glycine max aspartate kinase.
 Soy bean; aspartate kinase; cell transformation; transgenic plant.
 Glycine max.
 W0200146393-A2.
 28-JUN-2001.
 19-DEC-2000; 2000MO-US034396.
 21-DEC-1999; 99US-0172944P.
 (DUPO) DU PONT DE NEMOURS & CO E I.
 Falco SC, Farnodu CO, Thorpe CJ;
 WPI; 2001-418057/44.
 Novel nucleic acid molecules encoding aspartate kinase useful in cell transformation and transgenic plant production procedures.
 Example 3; Fig 1; 64pp; English.
 The present invention relates to nucleic acid molecules encoding aspartate kinase useful in cell transformation and transgenic plant production procedures. Nucleic acid encoding aspartate kinase is useful for producing a transgenic plant, by transforming a plant cell with a chimeric gene comprising nucleic acid encoding aspartate kinase operably linked to a regulatory sequence and regenerating a plant from the transformed plant cell. The present amino acid sequence is Glycine max aspartate kinase
 Sequence 564 AA;
 Query Match 65.2%; Score 1835; DB 4; Length 564;
 Best Local Similarity 70.6%; Pred. No. 2.3e-164;
 Matches 389; Conservative 42; Mismatches 80; Indels 40; Gaps 7;
 37 GPRGARGLSWVADSTSRRAKQADGGVLAGAPVL----- 71
 10 GVGQKLAVSMVSVRSLHCKSQI--GFALGAPVCARVWGNRAVSVTTCKASTDYIE 67
 72 -----GLGMEGLGDLQSVMTFGSSSVSSAARMAEVGLITPEERPVVLSAMGT 125
 68 KNALENGWVSSSE-ETSFCTVMKFGSSVASADRMKEVATLLSPPEEPVIVLSAMGT 126
 126 TNNLLAGEKAVGCGVIVHSEIEEMMVVSLIKTYDELGLP-XICNTSLYEGLLKGII 184
 127 TNNLLAGEKAVGCGVIVHSEIEELCPIDHILRTDQGVGVSISKLEELGLKGI 186
 185 AMMKELTPRTSDIVSSEGECSSTRISATLANKRVARROYDAFDGTTTDFEGNADILE 244
 187 AMMKELTPRTQDVLVSFGECMSTRIFAALINKGVARQYDAFEIGFITTDFTADILE 246
 245 ATYPAAVKRLHGMIOQPAIPVVTGFLGKMGSGAVTTIGRGSDLTATTICKALGRI 304
 247 ATPAAVKRLHGMISPAIATVTFGLGKARKSCAVTTIGRGSDLTATTICKALGRI 306
 305 QWMDVTCVITCPDNITPAKTYPIYLFEEATLAFGAQVTHPQSMRPARAGDIPVAVK 364
 307 QWMDVTCVITCPDNITPAKTYPIYLFEEATLAFGAQVTHPQSMRPARAGDIPVAVK 366
 365 NSYPAKPGTLITQRDMQGLVLTISIVLKSNTMLDIVSTRMLGOYFLARVSGICVI 424
 367 NSYPAKPGTLITQRDMQGLVLTISIVLKSNTMLDIVSTRMLGOYFLARVSGICVI 422
 425 EDLCISVDCAVTSEVSVSLDPSKIMRELIIQASLHDVVEELKAIYVLLQORAI 484

DB 423 EELGISVDVATSEVSVSLTDPKIMRELIIQASLHDVVEELKAIYVLLQORAI 482
 QY 485 SLIGNVEQSSLTILEKTGRVLRSGVNVGMISGASKVNSLIVHSDAKALVEALHQAFF 544
 DB 483 SLIGNVQSSLTILEKTGRVLRSGVNVGMISGASKVNSLIVHSDAKALVEALHQAFF 542
 QY 545 EDDVLSQVDAE 555
 DB 543 ESE-LSELEMD 552
 RESULT 4
 ABB93449
 ID ABB93449 standard; protein; 569 AA.
 AC ABB93449;
 DT 31-MAY-2002 (first entry)
 DE Herbicidally active polypeptide SEQ ID NO 2660.
 XX Herbicidal; plant; agriculture; herbicide.
 XX Arabidopsis thaliana.
 OS W0200210210-A2.
 XX 07-FEB-2002.
 XX 28-AUG-2001; 2001WO-EP009892.
 XX 28-AUG-2001; 2001WO-EP009892.
 XX (FARB) BAYER AG.
 PA Tietjen K, Weidler M;
 WPI; 2002-269010/31.
 DR Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms.
 XX Claim 5; SEQ ID NO 2660; 261pp + Sequence Listing; English.
 PS The invention relates to identifying target proteins (ABB90790-ABB94016)
 CC for herbicidally active compounds, comprising aligning and comparing
 CC nucleic acid or amino acid sequences from plant with nucleic acid or
 CC amino acid sequences from non-plant organisms using suitable search
 CC parameters, where plant sequences having an E-value greater by a factor
 CC of 3 than the E-value of most similar non-plant sequences are selected.
 CC The polypeptides or nucleic acids encoding them are useful for
 CC identifying modulators. The identified modulators are useful as
 CC herbicides
 SO Sequence 569 AA;
 Query Match 63.6%; Score 1789; DB 5; Length 569;
 Best Local Similarity 71.8%; Pred. No. 5.4e-160;
 Matches 369; Conservative 54; Mismatches 79; Indels 12; Gaps 6;
 44 LSWVWADSTSRRAKQADGGVLAGAPVLGGLGMEGLGJ-----QLSVWKFSGSSVSSAAR 99
 44 LSLPIDGSSITKVSAGSSGRNIVRA-VLEKKTETAITVDKGTICWKFSGSSVSSAAR 102
 QY 100 MAEVAQLITPEERPVVLSAMGKTNNLLAGEKAVGCGVIVHSEIEEMMVVSLIK 159
 DB 103 MKEVADLITPEERPVVLSAMGKTNNLLAGEKAVGCGVIVHSEIEEMMVVSLIK 162
 QY 160 TVDELGL-XICNTSLYEGLLKGIAAMKELTPRTSDIVSSEGECSSTRISATLANKR 218

[illegible]

	Query Match	Best Local Similarity	62.1%;	Score 1747.5;	DB 3;	Length 544;
	Matches 351;	Conservative 50;	Mismatches 65;	Indels 5;	Gaps 3	
Pr	25-OCT-1999;	99US-0161404.P				
Pr	25-OCT-1999;	99US-0161405.P				
Pr	25-OCT-1999;	99US-0161406.P				
Pr	26-OCT-1999;	99US-0161355.P				
Pr	26-OCT-1999;	99US-0161360.P				
Pr	26-OCT-1999;	99US-0161361.P				
Pr	26-OCT-1999;	99US-0161920.P				
Pr	28-OCT-1999;	99US-0161922.P				
Pr	28-OCT-1999;	99US-0161993.P				
Pr	29-OCT-1999;	99US-0162142.P				
Qy	78 GLGDQSYVMKGGSSVSSAARMAEVAGLLITFEERPVVVLISAMGKTINNLAGEKAV 137					
Dd	77 GSGKELTCVMKGGSSVSSAERMEKAVANLISPFDESPVILISMAGKTINKLKAGEKAV 136					
Qy	138 GCYVHVSEIIEEMNVKSLHKTVDLGL-PTXCNTSLYELBOLKGIAMMKELTRTS 196					
Dd	137 TCGVNVSEIEELSTFKELHRTAHEGVEITVLEKHEGLHQLKGISMKELTLTRD 196					
Qy	197 YLVSFEGCWSTRIFFAVYANKIRVKARQYDAFDIGFITTFEGNADILEATYPAVKRLHG 256					
Dd	197 YLVSFEGCSTHLPFAVYANKIGHARQYDAFEIGFITTDPTNADILEATYPAVSKTVG 256					
Qy	257 DWIQDPAIPVWTGFGKWKSGAVTTLGRGSDLTATTIGKALGRELQYWKVDGVLTC 316					
Dd	257 DWSKENAVVAVVGVYGGKWRSCAITTLGRGSDLTATTIGKALGRELQYWKVDGVLTC 316					
Qy	317 DPPIYVHAKTVYLLTFEEATELAVGAQVLIHPQSMRPARREGDIPVRVNSYNPKADGTLI 376					
Dd	317 DPPIYVGAQSVYLLTFDEAAELAVGAQVLIHPLSMRPARDDIPVRVNSYNPAPGTVI 376					
Qy	377 TRORDXALVLTSTVYLSKNTMLDIYSTRMLQGYGFLARVSGICIEELCISYDCAV 436					
Dd	377 TRERDNRK--AVLTIVYKENVMTMDIASTMLQGYGLAVY--FTTFEDLISVDVAT 432					
Qy	437 SEVSVSLSDPKSIKRSRELIOASELDHVELEKIAIVLLQORAIISLIGNVQSSLI 496					
Dd	433 SEVISILTLDPAKLGRRELIGRVNELDLVELEKIAVVKLLGRSIIISLIGNVQSSLI 492					
Qy	497 LKLTGVAKSGSVNTQMTSQGASKYMSLLIYHSDAKALYEAHQAFEND 547					
Dd	493 LKRVFVFSKNGVNVQMTSQGASKYMSLLIYHDEAEQCFALHSAFFED 543					
RESULT 6						
ABB93460	standard; protein; 544 AA.					
ID	ABB93460					
XX	ABB93460;					
XX	31-MAY-2002 (first entry)					
XX	Herbicideally active polypeptide SEQ ID NO 2671.					
XX	Herbicideal; plant; agriculture; herbicide.					
XX	Arabidopsis thaliana.					
XX	WO200210210-A2.					
XX	07-FEB-2002.					
XX	28-AUG-2001; 2001WO-EP009892.					
XX	28-AUG-2001; 2001WO-EP009892.					
XX	(FARB) BAYER AG.					
XX	Tietjen K, Weidler M;					
XX	11					

PR 28-JUN-1999; 99US-0140823P

[illegible]

QY	478	LOCRAIISLIGNVEOSSLLEKTGRVLRKSGVNVOMISQGSKRYMSLIVHSDPAKLVE	537
Db	493	LCORRSIISIGNQKSSLLERVFQVFRSNGVNVOMISQGSKRYNISLIVDEABOCVR	552
QY	538	ALHQAFEDD	547
Db	553	ALHSAFFED	562
RESULT 8			
AAAG46237	standard; protein; 459 AA.		
AAAG46237;			
AAAG46237;			
18-OCT-2000	(first entry)		
Arabidopsis thaliana protein fragment SEQ ID NO: 58149.			
Protein identification; signal transduction pathway; metabolic pathway;			
hybridisation assay; genetic mapping; gene expression control; promoter;			
termination sequence.			
Arabidopsis thaliana.			
EP1033405-A2.			
06-SEP-2000.			
25-FEB-2000; 2000EP-00301439.			
25-FEB-1999;	99US-0121825P.		
05-MAR-1999;	99US-0123180P.		
09-MAR-1999;	99US-0123548P.		
23-MAR-1999;	99US-0125788P.		
25-MAR-1999;	99US-0126264P.		
29-MAR-1999;	99US-0126785P.		
01-APR-1999;	99US-0127462P.		
06-APR-1999;	99US-0128234P.		
08-APR-1999;	99US-0128714P.		
16-APR-1999;	99US-0129645P.		
19-APR-1999;	99US-0130077P.		
21-APR-1999;	99US-0130449P.		
23-APR-1999;	99US-0130510P.		
23-APR-1999;	99US-0130891P.		
28-APR-1999;	99US-0131449P.		
30-APR-1999;	99US-0132048P.		
30-APR-1999;	99US-0132407P.		
04-MAY-1999;	99US-0132484P.		
05-MAY-1999;	99US-0132485P.		
06-MAY-1999;	99US-0132486P.		
06-MAY-1999;	99US-0132487P.		
07-MAY-1999;	99US-0132863P.		
11-MAY-1999;	99US-0134218P.		
14-MAY-1999;	99US-0134218P.		
14-MAY-1999;	99US-0134221P.		
14-MAY-1999;	99US-0134370P.		
18-MAY-1999;	99US-0134768P.		
19-MAY-1999;	99US-0134941P.		
20-MAY-1999;	99US-0135114P.		
21-MAY-1999;	99US-0135353P.		
24-MAY-1999;	99US-0135629P.		
25-MAY-1999;	99US-0136392P.		
27-MAY-1999;	99US-0136782P.		
28-MAY-1999;	99US-0137222P.		
01-JUN-1999;	99US-0137528P.		
03-JUN-1999;	99US-0137502P.		
04-JUN-1999;	99US-0137742P.		
07-JUN-1999;	99US-0138094P.		
08-JUN-1999;			

[illegible]

```

PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149922P.
PR 20-AUG-1999; 99US-0149902P.
PR 22-AUG-1999; 99US-0149930P.
PR 22-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151067P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 01-SEP-1999; 99US-0151308P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 15-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155599P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157653P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 16-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160982P.
PR 23-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161932P.
PR 28-OCT-1999; 99US-0161933P.
PR 28-OCT-1999; 99US-0162112P.
PR 29-OCT-1999; 99US-0162112P.

```

```

Query Match 61.4%; Score 1727.5; DB 3; Length 459;
Best Local Similarity 75.1%; Pred. No. 2.6e-154;
Matches 347; Conservative 48; Mismatches 62; Indels 5; Gaps 3;

```

87 MKFGSSVSSAARAAEVAAGLITFPERPVVIVLSANGKTTNILLAGEKAVGCGVTHSE 146

```

DB 1 MKFGSSVSSAARAAEVAAGLITFPERPVVIVLSANGKTTNILLAGEKAVGCGVTHSE 60
147 IEEEMNVXSLHIKIVDEIGL-PRICNTSLVEBOLKIGIAMKELTPTSDIYVSFECM 205
61 IEELSFIEKELHRIHAEHVEETVEIEKLEHQLKIGISMKELTTRTRDYVSFECM 120
206 STRIFSAYLNKIRVARQYDAFDIGFTTDEFGNADILEATYPAAVRLHGDWICDPAIP 265
121 STRIFSAYLNKIRVARQYDAFEIGFTTDEFTNADILEATYPAAVSKTIWGDMSKNAVP 180
266 VVTGFLGKMGKSGAVTTIGRGSDITATTIGKALGRBIOYKQVDCVLTCDPIYTHAK 325
181 VVTGYLGKMGKSCALITTTGSGSDITATTIGKALGRBIOYKQVDCVLTCDPIYTHAQ 240
326 TVPYLTFPEATLAFYQVIVHPQSMRAREGDIPIVKNKYNKARAGTLITRRDMDXG 385
241 SVPYLTFPEATLAFYQVIVHPQSMRAREGDIPIVKNKYNKARAGTLITRRDMDXG 299
386 LVLTSLVKSNTVMTLDIVSTRMLGOYFLARVSGICYIEDLCISVDCVATSEVSVSL 445
300 -AVLTSLVKSNTVMTLDIVSTRMLGOYFLAKV--FTTFEDLGISVDVATSEVSISSL 356
446 DPKTWSRRLIQASELDHVVBELEKIVALLQORALISLIGVQSSILTEKTVLR 505
357 DPAKLMGRRLIQVVELDNLVELEKIVALLQORALISLIGVQSSILTEKTVLR 416
506 KSGVNVQMSIGASKVNMSLIYHDSADALYEAHQAFEDD 547
417 SNGVNVQMSIGASKVNMSLIYHDSADALYEAHQAFEDD 458

RESULT 9
AAG46216
ID AAG46216 standard; protein; 478 AA.
XX
AC AAG46216;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58121.
XX
KW protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX
XX Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EB-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 23-MAR-1999; 99US-0126268P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 23-APR-1999; 99US-0131499P.
PR 28-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.

```

PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140355P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140823P.
PR 30-JUN-1999; 99US-0140823P.
PR 01-JUL-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142307P.
PR 09-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144844P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148313P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151430P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153738P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155466P.
PR 24-SEP-1999; 99US-0156458P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157783P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158223P.
PR 12-OCT-1999; 99US-0158363P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160747P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.


```

PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142802P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145152P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 26-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147320P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148655P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 25-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151067P.
PR 27-AUG-1999; 99US-0151068P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-015138P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.

PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155466P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156566P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159283P.
PR 13-OCT-1999; 99US-0159284P.
PR 13-OCT-1999; 99US-0159285P.
PR 14-OCT-1999; 99US-0159325P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160982P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161362P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 59.6%; Score 1675.5; DB 3; Length 446;
Best Local Similarity 74.8%; Pred. No. 2.1e-149;
Matches 336; Conservative 48; Mismatches 60; Indels 5; Gaps 3;

QY 100 MAEVAGLITPEERPVVLSAMKTTNNLLAGEKAVGGVTHVSEIEEMNVKSLHTR 159
DB 1 MKEVANLITSPDERPVVLSAMKTTNNLLKAGEKAVTGVNVSIEELSPKELHTR 60
QY 160 TYDEIGL-PIXCNLSYVLEBOLLKGIAMKELTPRTSDYVVSFGECMSTRIFSAVYINKTR 218
DB 61 TAEHGVETTVIEKHEBGLHQLGISMKELTIRTDYVVSFGECMSTRIFSAVYINKIG 120
QY 219 VKARQYDAFDIGFTITTEDEFGNADILEATYPAVAKRLHDMWIODPAIPVVTGFLGKWKSG 278
DB 121 HKARQYDAFELGFTITTDFTNADILEATYPAVSKTLVGDMSKENAVVTVGKGRSC 180
QY 279 ATTITGRCGSDLTATTIGKALGELIOWKDVGVCLCDNNTIPHAKTVRYLFFEEATEL 338
DB 181 ATTITGRCGSDLTATTIGKALGELIOWKDVGVCLCDNNTIPGASVRYLFFDEAAEL 240
QY 339 AYFGAQVLAPOSMPAREGDIPIVAVKSNYPKAPGTLITRQDMDXGLVLTJIVKSNV 398
DB 241 AYFGAQVLAHPSMPARAGDIPIVAVKSNYPKAPGTLITRQDMXGLVLTJIVKSNV 298
QY 399 TMLDVTSTRMGQGFHLARVSGICIEDICISVUCVATSEVSVSISDPEKISRELIQ 458
DB 299 TMLDVTSTRMGQGFHLARVSGICIEDICISVUCVATSEVSVSISDPEKISRELIQ 356
QY 459 ASELDHVAEELKLAIVLLOQRAIISLIGNVEOSSILKKTGVLKRSQVNTOMISOGA 518
DB 357 VNEIDNVIELEKLAIVVLLQRSIISLIGNVQSSILKKTGVLKRSQVNTOMISOGA 416

```

QY 519 SKYMSLIVHSDAKALVEALHQAFFEDD 547
DB 417 SKYMSLIVHDEEACQCRALHSAFFSTD 445

RESULT 11
AAG05075
ID AAG05075 standard; protein; 555 AA.

AC AAG05075;
XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 1342.

XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.

XX Arabidopsis thaliana.
OS EPI033405-A2.
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-00301439.
PF
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137523P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139482P.

PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 22-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 23-JUL-1999; 99US-0145276P.
PR 26-JUL-1999; 99US-0145513P.
PR 27-JUL-1999; 99US-0145516P.
PR 27-JUL-1999; 99US-0145618P.
PR 27-JUL-1999; 99US-0145619P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148584P.

PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149829P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151802P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151388P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 13-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 28-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160769P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 59.64; Score 1675.5; DB 3; Length 555;
Best Local Similarity 73.54; Pred. No. 2.9e-149;
Matches 347; Conservative 47; Mismatches 71; Indels 7; Gaps 4;

QY 82 QLSVVMKFGSSVSSAARMAEVAGLITPPEERPVVILSAMGKTNNLILAGEKAVCGV 141
DB 78 KLTCEVMKFGSSVSSAARMAEVAQLITPPEERPVVILSAMGKTNNLILAGEKAVCGV 137
QY 142 IHVSEIEEMNVKSLHKTVDLGL-PXICNTSLVLEBOLLGIMMKEITPRTSDYLS 200
DB 138 TIVDTIELSLITKELHRTIAHBLGVTAVIAEHLBOLLGIMMKEITPRTSDYLS 197

QY 201 FGEONSTIFSAVYANKIRVAKROYDADIGFITTTDERGNADILEATYPVAKRLHGDWIG 260
DB 198 FGEONSTIRLFAVYANKIRVAKROYDADIGFITTTDERGNADILEATYPVAKSKLLJDDWIK 257
QY 261 DPAIPVYVGTGFGKMGKGAVALTLGRGSDLTATIGALGLREIQVWKVDVGLTCDPNI 320
DB 256 ENALPVYVGTGFGKMGKSCAVTLTLGRGSDLTATIGALGLREIQVWKVDVGLTCDPNI 317
QY 321 YEHAKTVYLTFFBEATELAFGAQVLIHPQSMRPARBEDIYVRKNSYNPAFETLITROR 380
DB 318 YCGAQPVEHLTFDEAEELAFGAQVLIHPQSMRPARBEDIYVRKNSYNPAFETLITROR 377
QY 381 DMDXGLVLTIVKSNVTMLDIYSTMLGQVGLARVSGICIEDICISVDCVATSEVS 440
DB 378 DMSK--AVLTISLVKRNVTMLDIYSTMLGQVGLARV--FSTPEKLGISVDVATSEVS 433
QY 441 VVSVDPSKINSRELIQASSELHVEELEKIALVRLIQAPATISLIGNVQSSLIEXT 500
DB 434 ISLTLDPSKFCSEELIQH--ELDQVEELEKIALVNLIRHRSIISLIGNVQSSFILEKG 491
QY 501 GVLAKSGVNTQMTISQASKYNMELIVHSDPAKALVEALHQAPEDDVLSQV 552
DB 492 FVLKRTNGINQMTISQASKYNISLIVNDEALEHCVALHSAPEDDTCEAV 543

RESULT 12
AAG05074
ID AAG05074 strand: protein, 559 AA.
AC AAG05074;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 1341.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 1341.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
CS Arabidopsis thaliana.
XX
XX EP1033405-A2.
PN
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
PF
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129645P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130445P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132049P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135529P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139545P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143654P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145068P.
PR 21-JUL-1999; 99US-0145068P.
PR 22-JUL-1999; 99US-0145065P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145274P.
PR 26-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149925P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150565P.
PR 26-AUG-1999; 99US-0150864P.
PR 27-AUG-1999; 99US-0150658P.
PR 27-AUG-1999; 99US-0151068P.
PR 27-AUG-1999; 99US-0151082P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154739P.
PR 20-SEP-1999; 99US-0154799P.
PR 22-SEP-1999; 99US-0155138P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157153P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158332P.
PR 12-OCT-1999; 99US-0158363P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 13-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159584P.
PR 18-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.


```
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158332P.
PR 12-OCT-1999; 99US-0158393P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160983P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161599P.
PR 26-OCT-1999; 99US-0161606P.
PR 26-OCT-1999; 99US-0161611P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162143P.
```

Query Match 59.2%; Score 1666.5; DB 3; Length 473;
Best Local Similarity 73.3%; Pred. No. 1.6e-148;
Matches 345; Conservative 45; Mismatches 70; Indels 7; Gaps 4;

```
QY 87 MKFGSSVSSAAMAEVAGILTFPEERPVYVLSAMGXTNNLLAGEKAVGCVIHYSE 146
DB 1 MKFGSSVSAERKIVQAKILTFPEERPVYVLSAMAKTNNLLAGEKAVGCVIHYSE 60
QY 147 IEEWNVKSLHAKITVDELGI-PXICNTSLYEISOLKSIAMKELTPRTSDYLVFGEQM 205
DB 61 IEEISYIKELHIRTAHLEIVETAVIAEHLEGLQLKGVAMKELTLRSRDYLVFGEQM 120
QY 206 STRFSAVYLNKIRKAKQVADAFIGFTTDFEGNADILEATYPAVAKRLHGMICDPAIP 265
DB 121 STRFPAVYLNKIRKAKQVADAFIGFTTDFEGNADILEATYPAVAKRLHGMICDPAIP 180
QY 266 VVTGFLGKWKSGAVITLGRGSDLTATTIGKALGREIQWEDVDGVLTCDPNTYCGAQ 325
DB 181 VVTGFLGKWRSCAVITLGRGSDLTATTIGKALGREIQWEDVDGVLTCDPNTYCGAQ 240
QY 326 TVPILTFEEATELAYPAQVILHPSMPPARESDIPRVKNSVYPKAGFLITRODMDXG 385
DB 241 PVPHLTDEALELAVYPAQVILHPSMPPARESDIPRVKNSVYPTLPGVITRSDMGK- 299
QY 386 LVVLTSLVKSNTMDIVSTRMLGQYGLFARVSGICYIEDLCISVDCVATSEVSVSL 445
DB 300 -AVLTSLVKNVMTLDTSTRMLGQYGLFARV--FSTEKLGISVDVATSEVSIISLT 356
QY 446 DPKIWSRELIIQASLIDHVEELKTAIVRLIQARLISLGNVQSSLIIEKRGVYR 505
DB 357 DPKFCSRELIIQ--ELDQVEELKTAIVNLLHRSIISLISGNVQSSFIIEKRGVYR 414
```

```
QY 506 KSGVNVQMSIQASKVMNSLIYHSDAKALVEALHQAFFEDPYLSQY 552
DB 415 TNGINVQMSIQASKVMNSLIYHSDAKALVEALHQAFFEDPYLSQY 461

RESULT 15
AA646217 standard; protein; 465 AA.
AC AA646217;
XX 18-OCT-2000 (first entry)
DB Arabidopsis thaliana protein fragment SEQ ID NO: 58122.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridization assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0123788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0128845P.
XX 18-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132863P.
XX 14-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
```


Db 61 VHSPLALFSEWEIKFSMERTAEHGLVETTVEIEKLEGLHQLKGISMKKELLTRNDYLV 120
QY 200 SFGECMSTRIFSAVYLNKIRVAKROYDAFDIGFITTDEFGNADILEATYPVAXRLHQDWI 259
Db 121 SFGECMSTRIFSAVYLNKIGHKARQYDAFIEIGFITTDDFTNADILEATYPVASKTVGDWS 180
QY 260 ODPALPVVYGLFGKMGKSGAVTTLGRGSSDLJATTIGKALGREIOVMKDVGVLTCDPN 319
Db 181 KENAVPVVYGLFGKMGKSCAITTLGRGSSDLTATTIGKALGREIOVMKDVGVLTCDPN 240
QY 320 IYPAKTVPYLTFEEATELAYFGAQLVPCQMPARBGDIPIRYKNSYNPKAPGTLITRQ 379
Db 241 IYPAQSVPYLTFDEAAELAYFGAQLVPLSRPARBGDIPIRYKNSYNPKAPGTVITRS 300
QY 380 RMDXGLVYLTSLVLSKNTMLDIVSTRMLGQYGFARVSGICYIEDLCISVDCVATSEV 439
Db 301 RDMSK--AVLTSLVLSKNTMLDIASSTRMLGQYGFARV--FTTFEDLGISVDVATSEV 356
QY 440 SVSVSLDPKXMSRELIOQASFLDVVELEKIALVRLQORATISLIQNVQSSLILEK 439
Db 357 SISLTLDPAKLMGRELIOQVNELDNLVELEKIAVAVKLLORRSIISLIQNVQSSLILEK 416
QY 500 TGSVLKSGVNTOMISOGASKVMSLIYHSDPAKALVEALHQAFFEDD 547
Db 417 VFOVFRSNGVNTOMISOGASKVMSLIYHSDPAKALVEALHQAFFEDD 464

Search completed: March 23, 2004, 13:15:47
Job time : 85 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 13:14:31 ; Search time 24 Seconds
(without alignments)
1204.606 Million cell updates/sec

Title: US-09-890-813-6

Perfect score: 2813

Sequence: 1 MALPVSAAAPRLVPSIPP.....QAFEDVLGVYEANLLVG 560

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfilest.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	590	21.0	453	US-09-489-039A-8611	Sequence 8611, Ap
2	569	20.2	449	US-08-256-136-2	Sequence 2, Appl
3	569	20.2	449	US-08-950-737-2	Sequence 2, Appl
4	569	20.2	449	US-08-973-461A-8	Sequence 5, Appl
5	569	20.2	449	US-08-648-010-8	Sequence 8, Appl
6	565	20.1	460	US-09-543-681A-7248	Sequence 7248, Ap
7	469.5	16.7	835	US-09-489-039A-8740	Sequence 8740, Ap
8	469.5	16.5	820	US-08-380-182-23	Sequence 23, Appl
9	455.5	16.2	862	US-08-543-681A-6315	Sequence 6315, Ap
10	428.5	15.2	478	US-09-107-532A-6030	Sequence 6030, Ap
11	416.5	14.8	863	US-08-380-182-19	Sequence 19, Appl
12	415.5	14.8	463	US-09-134-001C-4160	Sequence 4160, Ap
13	414.5	14.7	421	US-08-532-828B-3	Sequence 3, Appl
14	414.5	14.7	421	US-08-532-828B-4	Sequence 4, Appl
15	414.5	14.7	421	US-08-700-359-9	Sequence 9, Appl
16	414.5	14.7	421	US-08-596-366-6	Sequence 6, Appl
17	414.5	14.7	421	US-08-967-104-6	Sequence 6, Appl
18	414.5	14.7	421	US-08-985-908-5	Sequence 5, Appl
19	414.5	14.7	421	US-08-853-730-14	Sequence 14, Appl
20	414.5	14.7	421	US-08-985-916-5	Sequence 5, Appl
21	414.5	14.7	863	US-08-380-182-20	Sequence 20, Appl
22	402	14.3	411	US-07-684-135A-2	Sequence 2, Appl
23	402	14.3	421	US-08-311-731A-30	Sequence 30, Appl
24	377.5	13.4	461	US-09-252-991A-29720	Sequence 29720, A
25	366.5	13.0	438	US-09-328-352-5111	Sequence 5111, Ap
26	366.5	13.0	828	US-09-489-039A-12219	Sequence 12219, A
27	334.5	11.9	441	US-09-198-452A-1124	Sequence 1124, Ap

28	325	11.6	424	US-09-134-001C-5225	Sequence 5225, Ap
29	300	10.7	815	US-09-543-681A-8330	Sequence 8330, Ap
30	294.5	10.5	285	US-09-540-236-3031	Sequence 3031, Ap
31	246	8.7	256	US-09-134-000C-6771	Sequence 6771, Ap
32	196	7.0	262	US-08-687-534A-23	Sequence 23, Appl
33	195	7.0	262	US-09-527-431-23	Sequence 23, Appl
34	187.5	6.7	156	US-09-056-556-227	Sequence 227, App
35	187.5	6.7	156	US-09-072-596-227	Sequence 227, App
36	187.5	6.7	156	US-09-072-967-227	Sequence 227, App
37	116	4.1	528	US-09-543-681A-6385	Sequence 6385, Ap
38	108.5	3.9	250	US-09-543-681A-6626	Sequence 6626, Ap
39	105.5	3.8	432	US-09-252-991A-18147	Sequence 18147, A
40	105	3.7	1705	US-08-669-785-4	Sequence 4, Appl
41	104.5	3.7	702	US-09-328-352-8176	Sequence 8176, Ap
42	104	3.7	469	US-08-985-335-5	Sequence 5, Appl
43	102.5	3.6	528	US-09-410-372-5	Sequence 5, Appl
44	102.5	3.6	528	US-09-134-001C-4262	Sequence 4262, Ap
45	102.5	3.6	582	US-09-252-991A-25366	Sequence 25366, A

ALIGNMENTS

RESULT 1
US-09-489-039A-8611
Sequence 8611, Application US/09489039A
Patent No. 6610816

GENERAL INFORMATION:
APPLICANT: Gary Breton et. al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-25
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 8611
LENGTH: 453
TYPE: PRT
ORGANISM: Klebsiella pneumoniae

US-09-489-039A-8611

Query Match 21.0%; Score 590; DB 4; Length 453;

Best Local Similarity 34.9%; Pred. No. 9, 2e-51;
Matches 166; Conservative 81; Mismatches 184; Indels 44; Gaps 15;

QY	85	VVMKFGGSSVSAARMAEVALITLTPERPVLVLSAMGKTNNLL-LAGEKAVCGVTH	143
DB	9	VVAKEGTSVADPDANRSIDVALLDANTR-IVLSASAGVNIIVALLAG-----GLEP	61
QY	144	VSEIEENWVKSLHKTVDLGLPXICNTSLVELBOLKGMKELPRIS-----DYL	198
DB	62	TERSCQDALRQIOFNILRLRPSVIRE---EIERLINITTLAEAMALASTALTDEL	118
QY	159	VSPEECSTRIEFAVINKIRVKROYDARDIGITDEGND-----TLKTPAVAKR	253
DB	119	VSHLELSTLTFEILRERICIQWPDARKV-LRTDRCGRAPDPAIAELTQOQLAR	177
QY	254	LHGWDIDPAIPVVTGFLGKWSGAVTTLGRGSDLVATTGKALGREIVMKVDGV	313
DB	178	L-----AEGVLTGIGSEAK-GRITTLGRGSDYTAALGELNATRVIMWDVGI	230
QY	314	LTCPNTPYPAKTVPIVTFEEATLVEFGAIVHPOSMKPARSGDIPVVKSYNPKAG	373
DB	231	YTPDPAAPAKRIDVIAFEAAEMATFGAKVHPATLPAVRSIDIPVVGSKPKAG	290
QY	374	TLITQRDMXGLVLTSLVLSKSNVTMLDIVSTRMLGOYGFARVSGICYEDLCISYDC	433
DB	291	TLVKTKE---NPLFPALALRRQTLTLTSLNMLHSGFLAEVVGILARHN--ISVDL	345
QY	434	VATSEVSVSLDPSKTSNR--ELIQASLHVVELEKIAIVLLOQRALISLIGN-V	490

Db 346 TTSEVSVALTMDTGTSGADTLLTQA-----LITELSLCRVESENLATVALIGNEL 400

QY 491 ESSLILEKTRGVLRKSGVNVOMISQASKYNMGLIYHSDPAKLYALHQAPE 545

Db 401 SRAQGVKEVFGVLEP--FNIRMLCYGASHNLGFLVPGDAEKVQKLNHNLE 453

RESULT 2

US-08-256-136-2

/ Sequence 2, Application US/08256136

/ Patent No. 5661012

/ GENERAL INFORMATION:

/ APPLICANT: SANO, KONOSUKE

/ APPLICANT: KOJIMA, HIROYUKI

/ APPLICANT: OGAMA, YURI

/ TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF L-THREONINE

/ TITLE OF INVENTION: BY FERMENTATION

/ NUMBER OF SEQUENCES: 4

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT

/ STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400

/ CITY: ARLINGTON

/ STATE: VIRGINIA

/ COUNTRY: USA

/ ZIP: 22202

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patent in Release #1.0, Version #1.30

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/256,136

/ FILING DATE: 01-JUL-1994

/ CLASSIFICATION: 435

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: PCT/JP93/01640

/ FILING DATE: 10-NOV-1993

/ PRIOR APPLICATION DATA: JP 300021/92

/ FILING DATE: 10-NOV-1992

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (703) 413-3000

/ TELEFAX: (703) 413-2220

/ TELEX: 248955 OPAT UR

/ INFORMATION FOR SEQ ID NO: 2:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 449 amino acids

/ TYPE: amino acid

/ TOPOLOGY: linear

/ MOLECULE TYPE: protein

US-08-256-136-2

Query Match 20.2%; Score 569; DB 1; Length 449;

Best Local Similarity 34.0%; Pred. No. 1.2e-48;

Matches 161; Conservative 89; Mismatches 182; Indels 42; Gaps 14;

QY 85 VVMKFGSSVSSAAAEVAGLITPPEERPVVLSAMGKTNNLLAGEKAVCGVIVH 144

Db 5 VVSKFGSTVADPDANRSADIVLSDANVR-LVLSASAGITNLVALAE-----GLEPG 58

QY 145 SEIEEMNVKSLHIKIVDELGLPKICNTSLYEQLKGIAMKE-----LTPRTSDIV 199

Db 59 EREKLDARNIQFALLERLRYP--NVIREIERLENTIVLAAMALATSPALTDELV 115

QY 200 SPEECMSTRIFSAYLKIRKARQAYDADIGITITDEFQND-----LLEATYPAVAKRL 254

Db 116 SHELMSTLLFVEILBERVOAQWPFVRKV-WRTNDRFGAPDPAALAEALALQLPRL 174

QY 255 HGMWIDPALPVVTGELGKMGKSGAVTTIGRGSGLIKATTIKALGLREIQWMDVQVTL 314

Db 175 -----NEGIVITQIGISENK-GRITTLGRGGSYTAALAEALHASRVIMIDVPGIY 227

QY 315 TCDPNIVPAKIVPVYLTPEATELAFGAQVHLQSPMPAREGDIPIRVKNSYVPAKPGT 374

Db 228 TTDPVRSAAKRIDEIAFAAEAMATFGALVLPATLLPAYSDDIPFVSGSKDPAGT 287

QY 375 LITQRMKXGLVYLTISIVKSNVTMDIVSTFMGLQGYFLARVSGICVIEDLCISYDCV 434

Db 288 LVCKRTE--NPFLRALRNRQTLTLHSLNMLHSRGLAEVFGILARHN--ISVDLI 342

QY 435 ATSEVSVALDPSKISR--ELIQASELDHYVEELEKAIYVRLIQSAIISLIGN-VE 491

Db 343 TTSEVSVALTMDTGTSGADTLLTQA-----LITELSLCRVESENLATVALIGNEL 397

QY 452 QSSLILEKTRGVLRKSGVNVOMISQASKYNMGLIYHSDPAKLYALHQAPE 545

Db 398 KACGVKEVFGVLEP--FNIRMLCYGASHNLGFLVPGDAEKVQKLNHNLE 449

RESULT 3

US-08-950-737-2

/ Sequence 2, Application US/08950737

/ Patent No. 5932453

/ GENERAL INFORMATION:

/ APPLICANT: KIKUCHI, YOSHIMI

/ APPLICANT: NAKAMISHI, KAZUO

/ APPLICANT: KOJIMA, HIROYUKI

/ TITLE OF INVENTION: PROCESS FOR PRODUCING L-AMINO ACID

/ TITLE OF INVENTION: THROUGH FERMENTATION

/ NUMBER OF SEQUENCES: 8

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,

/ STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400

/ CITY: ARLINGTON

/ STATE: VA

/ COUNTRY: USA

/ ZIP: 22202

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patent in Release #1.0, Version #1.30

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/950,737

/ FILING DATE: 15-OCT-1997

/ CLASSIFICATION:

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: JP 272114/1996

/ FILING DATE: 15-OCT-1996

/ ATTORNEY/AGENT INFORMATION:

/ NAME: OBLON, NORMAN F.

/ REGISTRATION NUMBER: 24,618

/ REFERENCE/DOCKET NUMBER: 10-888-0

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (703) 413-3000

/ TELEFAX: (703) 413-2220

/ INFORMATION FOR SEQ ID NO: 2:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 449 amino acids

/ TYPE: amino acid

/ TOPOLOGY: linear

/ MOLECULE TYPE: Protein

US-08-950-737-2

Query Match 20.2%; Score 569; DB 2; Length 449;

Best Local Similarity 34.0%; Pred. No. 1.2e-48;

Matches 161; Conservative 89; Mismatches 182; Indels 42; Gaps 14;

QY 85 VVMKFGSSVSSAAAEVAGLITPPEERPVVLSAMGKTNNLLAGEKAVCGVIVH 144

Db 5 VVSKFGSTVADPDANRSADIVLSDANVR-LVLSASAGITNLVALAE-----GLEPG 58

QY 145 SEIEEMNVKSLHIKIVDELGLPKICNTSLYEQLKGIAMKE-----LTPRTSDIV 199

Db 59 EREKLDARNIQFALLERLRYP--NVIREIERLENTIVLAAMALATSPALTDELV 115

QY 200 SGECSMTRIFSAIYANKRKYARQYDAFDIGFITTDEFGAND-----ILEATYPAVAKRL 254
 DB 116 SHELMSTILLFVEILIREDDVQAWFDRKV-WRTNDRFGRAEPDIAALAEALQILPRL 174
 QY 255 HGDWIODPAIPVVTGLGKMGSGAVTTLGRGSSDLTATTIGKALGRIQWMDYDGV 314
 DB 175 -----NEGVLITQGFISENK-GRITTLGRGSDYTTALLAEALHARSVDIWDVGIY 227
 QY 315 TCDENIYPAKTVPYLTPEATELAYFGAQLHQSMPAREGDIYRVKNSYNPKAPGT 374
 DB 228 TTDPRVSAKRIDEIAFAEAEMATFGAKVLPATLLPAYRSDIPFVGSXDPKRAAGT 287
 QY 375 LITRQDMOXGLVLTSTVLKSNVTMLDVSTRMLGQYGFARVSGICYIEDLCISYDCV 434
 DB 288 LVGNKTE--NPELFRALALRNQTLTLHSLNMLHSGFLAEVFGILARNH--ISVDLI 342
 QY 435 ATSEVSVSALDPSKIMSR--ELIQASELDHVEELEKIAIVRLQORAIISLIGN-VE 491
 DB 343 TISEVVALTLDTGSTGTGDTLLTQS-----LIMELSLACREVEEGALVALIGNDLS 397
 QY 492 QSLILEKTRGVLRKSGVNVQMSQASAKVMSLIYHSDAKALVEALHQAFFE 545
 DB 398 KACGVKEVFGVLEP--FNIRMICYGASSHNLCLVPGEDAEQVQKLSNLF 449

RESULT 4
 US-08-973-461A-8
 Sequence 8, Application US/08973461A

GENERAL INFORMATION:
 APPLICANT: KOJIMA, HIROYUKI
 APPLICANT: OGAWA, YURI
 APPLICANT: KANAMURA, KAZUE
 APPLICANT: SANO, KONOSUKE
 TITLE OF INVENTION: METHOD OF L-LYSINE BY FERMENTATION
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT, P.C.
 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
 CITY: ARLINGTON
 STATE: VA
 COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/973,461A
 FILING DATE: 20-APR-1998
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 7-146054
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: NORMAN F. OBLON
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 0010-0901-0 PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 449
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-973-461A-8

Query Match 20.2%, Score 569, DB 2, Length 449,
 Best Local Similarity 34.0%, Pred. No. 1,2e-48;

Matches 161; Conservative 89; Mismatches 182; Indels 42; Gaps 14;
 QY 85 VYMKGGSSVSSAARMAEVAGLITFPBERVVTLSMGKTTNNLLAGEKAVGCGYIV 144
 DB 5 VVSKFGGTSVADFDAMNSADIVLS DANVR-LVLLSASAGITNLLVLAEE-----GLEFG 58
 QY 145 SEIEEMNWKSLIHKTYDELGLPYICNTSLYEQLKIGIAMKE-----LTPRTSDYLV 159
 DB 59 ERFELDAIRNITQFALIRLRYF--NVIREFIRLLENITVLAEEAALATSPALTDLV 115
 QY 200 SGECSMTRIFSAIYANKRKYARQYDAFDIGFITTDEFGAND-----ILEATYPAVAKRL 254
 DB 116 SHELMSTILLFVEILIREDDVQAWFDRKV-WRTNDRFGRAEPDIAALAEALQILPRL 174
 QY 255 HGDWIODPAIPVVTGLGKMGSGAVTTLGRGSSDLTATTIGKALGRIQWMDYDGV 314
 DB 175 -----NEGVLITQGFISENK-GRITTLGRGSDYTTALLAEALHARSVDIWDVGIY 227
 QY 315 TCDENIYPAKTVPYLTPEATELAYFGAQLHQSMPAREGDIYRVKNSYNPKAPGT 374
 DB 228 TTDPRVSAKRIDEIAFAEAEMATFGAKVLPATLLPAYRSDIPFVGSXDPKRAAGT 287
 QY 375 LITRQDMOXGLVLTSTVLKSNVTMLDVSTRMLGQYGFARVSGICYIEDLCISYDCV 434
 DB 288 LVGNKTE--NPELFRALALRNQTLTLHSLNMLHSGFLAEVFGILARNH--ISVDLI 342
 QY 435 ATSEVSVSALDPSKIMSR--ELIQASELDHVEELEKIAIVRLQORAIISLIGN-VE 491
 DB 343 TISEVVALTLDTGSTGTGDTLLTQS-----LIMELSLACREVEEGALVALIGNDLS 397
 QY 492 QSLILEKTRGVLRKSGVNVQMSQASAKVMSLIYHSDAKALVEALHQAFFE 545
 DB 398 KACGVKEVFGVLEP--FNIRMICYGASSHNLCLVPGEDAEQVQKLSNLF 449

RESULT 5

US-08-648-010-8

Sequence 8, Application US/08648010
 Patent No. 6040160
 GENERAL INFORMATION:
 APPLICANT: KOJIMA, HIROYUKI
 APPLICANT: OGAWA, YURI
 APPLICANT: KANAMURA, KAZUE
 APPLICANT: SANO, KONOSUKE
 TITLE OF INVENTION: METHOD OF PRODUCING L-LYSINE BY
 FERMENTATION
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
 STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
 CITY: ARLINGTON
 STATE: VA
 COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/648,010
 FILING DATE: 29-MAY-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/JP94/01994
 FILING DATE: 28-NOV-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/JP94/01994
 FILING DATE: 28-NOV-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 5/308397
 FILING DATE: 08-DEC-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 10-805-0 PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 449 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-648-010-8

Query Match 20.2%; Score 569; DB 3; Length 449;
 Best Local Similarity 34.0%; Pred. No. 1,2e-48;
 Matches 161; Conservative 89; Mismatches 182; Indels 42; Gaps 14;

QY 85 VMKFGSSVSSAARMAEAVAGILTFPEERPVVLSAMGKTTNNLLAGEKAVGCGVIHV 144
 DB 5 VVSKFGTGVADPDMNRSADIVLSDANVR-LVYLSASAGITNLLVALAE-----GLEPG 58
 QY 145 SEIEEMNVKSLHIKTVDLGPXICNTSLVELBOLLKGIAMKE-----LTPRTSDYLV 139
 DB 59 EPEKLDALRNIGFALIERLRYF--NVIREIERLENITVLEMAALATSPALTDLVL 115
 QY 200 SPGEQSTRIFSAVYLNKIRVKARQYDAFDIGFITDEFGND-----ILEATYPAVAKRL 254
 DB 116 SHGELNSTLTFEILREKRGVQWQDFVAKV-NRTNDRGRABPDIAALAEIALQLPRL 174
 QY 255 HGDWIDDPALPVVTGFLGKMGKSGAVTTLGRGSDLTATTGKALGREIQWQDVGV 314
 DB 175 -----NEGVLITGGIGSENR-GRITTLGRGSDYVTAALAEALHSRVDIWTVDVGIT 227
 QY 315 TCDPNLYPAKTVPIYLFEEATELAYFGAQLHPQSMRPARBGDIPIRVKNSYNPKAGT 374
 DB 228 TTDPRVSAKRIIDELAFEAEMATFGAKVLHPATLLPAVRSDIPVVGSKCPRAAGT 287
 QY 375 LTRQDMXGLVLTSLVLSKSNVTMLDVSTRMLGOYGFARVSGICYIEDLCISVDCV 434
 DB 288 LVCNKTE--NPFLFALALRNQTLTLHSLNMLHSGFLAEFGIARHN--ISVDLI 342
 QY 435 ATSEVSVSLDPSKIWSR--ELIQASELDHVELEKIALVRLQORALISLIGN-VE 491
 DB 343 TISEVSVALTIDTGTSTGDTLLQS-----LMLELALCRVEVBEGIALVALIGDLS 397
 QY 492 QSSLLIEKTGRVLRKSGVNVQWISQASKNMSLIVHSDAKALVEALHQAFFE 545
 DB 398 KACGVKEVGVLEP--FNIRMIQYGASHNLCLFVPGDAEVOVQKLSNLF 449

RESULT 6
 US-09-543-681A-7248
 Sequence 7248, Application US/09543681A
 Patent No. 6605709
 GENERAL INFORMATION:
 APPLICANT: GARY BRETTON
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 2709, 1002-001
 CURRENT APPLICATION NUMBER: US/09/543, 681A
 CURRENT FILING DATE: 2000-04-05
 PRIOR APPLICATION NUMBER: US 60/128, 706
 PRIOR FILING DATE: 1999-04-09
 NUMBER OF SEQ ID NOS: 8344
 SEQ ID NO 7248
 LENGTH: 460
 TYPE: PRT
 ORGANISM: Proteus mirabilis
 US-09-543-681A-7248

Query Match 20.1%; Score 565; DB 4; Length 460;
 Best Local Similarity 34.3%; Pred. No. 3.2e-48;
 Matches 163; Conservative 82; Mismatches 182; Indels 48; Gaps 15;

QY 86 VMKFGSSVSSAARMAEAVAGILTFPEERPVVLSAMGKTTNNLLAGEKAVGCGVIHV 145
 DB 18 IAKFGGTSVANFAAMEKADIIKQSVR-VVYLSASAGITNLLI-----ELATVEARQ 71
 QY 146 EIEEMNVKSLHIKTVDLGPXICNTSLVELBOLLKGIAMKE-----LTPRTSDYLV 200
 DB 72 RAALLAQVDRLEVALINQISQAEIIS--QENIRLENIENTEMSEAAALATSPALTDLVL 128
 QY 201 FGEQSTRIFSAVYLNKIRVKARQYDAFDIGFI--TTDEFGND-----LLEATYPAVAR 253
 DB 129 HGEIINSTLTF--VELIREKGISADWFDVAKVKTNDLFCRAPDMAQLTELTQSIQCR 185
 QY 254 LHGDWIDDPALPVVTGFLGKMGKSGAVTTLGRGSDLTATTGKALGREIQWQDVGV 313
 DB 186 L-----TETVTVQSGTIGQEPK-GKITTLGRGSDYVTAALIEALGMSRVDIWTVDVGI 238

QY 314 LTRQDMXGLVLTSLVLSKSNVTMLDVSTRMLGOYGFARVSGICYIEDLCISVDC 433
 DB 239 YSTDPRVYPOQRIDHIAFDPAEMATFGAKVLHPATLLPAVRSGIPVVGSKAPBEG 298
 QY 434 VATSEVSVSLDPSKIWSR--ELIQASELDHVELEKIALVRLQORALISLIGN-- 489
 DB 354 ITTSVSTALTLDTGTSTSSGSLITNA-----LTELNALCRVEEEDLAVAILGNL 408
 QY 490 VQSSLLIEKTGRVLRKSGVNVQWISQASKNMSLIVHSDAKALVEALHQAFF 544
 DB 409 SQVNLGSGITGTL---EDNIRMIHSGASTHNLCLVQDKADNIVRKLDHLLF 460

RESULT 7
 US-09-489-039A-8740
 Sequence 8740, Application US/09489039A
 Patent No. 6610836
 GENERAL INFORMATION:
 APPLICANT: Gary Bretton et. al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 2709, 2004001
 CURRENT APPLICATION NUMBER: US/09/489, 039A
 CURRENT FILING DATE: 2000-01-27
 PRIOR APPLICATION NUMBER: US 60/117, 747
 PRIOR FILING DATE: 1999-01-29
 NUMBER OF SEQ ID NOS: 14342
 SEQ ID NO 8740
 LENGTH: 835
 TYPE: PRT
 ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-8740

Query Match 16.7%; Score 469.5; DB 4; Length 835;
 Best Local Similarity 29.1%; Pred. No. 3.8e-38;
 Matches 141; Conservative 101; Mismatches 207; Indels 35; Gaps 13;

QY 86 VMKFGSSVSSAARMAEAVAGILTFPEERPVVLSAMGKTTNNLLAGEKAVGCGVIHV 144
 DB 18 VIKFGGTSVANFAARERLRADILHESNAROGVATVLSAPATITNHLVMIKIGTIG-QGDAL 76
 QY 145 SEIEEMNVKSLHIKTVDL-----LGLPYICNTSLVELBOLLKGIAMKEELTPRTSDY 197
 DB 77 PNIAAERIFALLQGLADAPAPPLAQLKAFVQBEFAQIKVHLHGLSLGQCPDSVNA 136
 QY 198 LVSFGQSTRIFSAYL-----NKIRVKARQYDAFDIGFITDEFGNDILEATYPAVAR 253
 DB 137 LICRGKSLIALMAGLBARGHKVSVINPVKLAAGHYLE---STYDIAESRRLAAAG 193
 QY 254 LHGDWIDDPALPVVTGFLGKMGKSGAVTTLGRGSDLTATTGKALGREIQWQDVGV 313
 DB 194 IPAD-----HMTIMAGFTA-GNEKGEIVLVLRGNSDYSAVLAACLRADCCETIWTVDGV 247

Db 219 ALNIPKDHMLMAGFTA-GNEKNELVILGRNGSDYSAVAALACLAQCCEIWTVDGYT 277
 QY 316 CDPNITPAKTVPYLTFFEEATLAFGAQVLPQSMRPARBEDIPIVRKNSYNKAGTL 375
 Db 278 CDRILVPDAHLKGRSFGAMELSYFGAKVLPRTIAPACQIPLCKTNGPEAFGLT 337
 QY 376 I-TRORDMXGLVLTSLIVLKSNTWMLDIVSTRMLGOYGFPLARVSGICIEDLCISYDCV 434
 Db 338 IGGQKDDSTPVYGIINL---NNMAMINVSQPMGMGMVGMARVPSVMSRAGISVLTQ 394
 QY 435 ATSEVSVSLSDSKTSRELIQASL-----DHVELEKIAIVLLQORAIIS 485
 Db 395 SSSEVSISPCV-PQK---ELIRAKALSEEFYLLKOGVLDPLD-----IMNNVAIIS 443
 QY 486 LIQN-VEQSLILEKTGRVLRKSGVNVQMSOGASKVMSLIVHSDAKALVEALHQAFF 544
 Db 444 VVDGMRITLKGIRARFPAITRGNITVAIAQSSERSISAVIANDATNVRLOHQLF 503
 QY 545 EDDVLSQV 552
 Db 504 NTQIVEV 511
 RESULT 10
 US-09-107-532A-6090
 / Sequence 6090, Application US/09107532A
 / Patent No. 6583275
 / GENERAL INFORMATION:
 / APPLICANT: Lynn A Doucette-Stamm and David Bush
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 / ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 / NUMBER OF SEQUENCES: 7310
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 / STREET: 100 Beaver Street
 / City: Waltham
 / STATE: Massachusetts
 / COUNTRY: USA
 / ZIP: 02354
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: CD-ROM ISO9660
 / COMPUTER: PC
 / OPERATING SYSTEM: <Unknown>
 / SOFTWARE: ASCII
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/09/107,532A
 / FILING DATE: 30-Jun-1998
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: 60/085,598
 / FILING DATE: 14 May 1998
 / APPLICATION NUMBER: 60/051571
 / FILING DATE: July 2, 1997
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Arinello, Pamela Deneka
 / REGISTRATION NUMBER: 40,489
 / REFERENCE/DOCKET NUMBER: GTC-012
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (781)893-5007
 / TELEFAX: (781)893-8277
 / INFORMATION FOR SEQ ID NO: 6090:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 478 amino acids
 / TYPE: amino acid
 / TOPOLOGY: linear
 / MOLECULE TYPE: protein
 / HYPOTHEICAL: YES
 / ORIGINAL SOURCE:
 / ORGANISM: Enterococcus faecium
 / FEATURE:
 / NAME/KEY: misc_feature
 / LOCATION: (B) LOCATION 1...478
 / SEQUENCE DESCRIPTION: SEQ ID NO: 6090:
 US-09-107-532A-6090

Query Match 15.2%; Score 428.5; DB 4; Length 478;
 Best Local Similarity 27.2%; Pred. No. 2.1e-34;
 Matches 130; Conservative 96; Mismatches 193; Indels 59; Gaps 15;
 QY 86 VMKFGSSVSSAARMAEYAGLITFFPEERPVVLSAMGKTNNLLAGEAVACGYI--- 142
 Db 32 VAKFGTSMSEHQFKVRRITVLSDP-NRQIVVSAIGKRTSD---DDKVTDLVLYIA 86
 QY 143 -----HVEIEEMNMY--KSLIKTVDELGLPXICNTSLVELEOLLKGIAMMKELTPRTSD 196
 Db 87 YLGHVDMTPMLKRICDFIQRVYALQLP-----IDEELARIQQLDEKITED 137
 QY 197 YLVSPFGCMSTRIFSAYINKIRVARQYDAFDIGFITTFEGNADILEA--TPYAVAKL 254
 Db 138 YLVSRGEVLTALQMAEVLGYQFIDAKDLIFPDYDGKIDQEKTAQELTKAFKRPKM---- 193
 QY 255 HGDWTDPAIPVYTGFLGKWSGAVTTLGGSGDLTATTIGKALGLREIOWKDYDGYL 314
 Db 194 -----VPGFYGAN-PAGEKILGRSGSDITGALITLQAESYENMTDVSQIM 241
 QY 315 TCDPNIYPKATVPYLTFFEEATLAFGAQVLPQSMRPARBEDIPIVRKNSYNPAPGT 374
 Db 242 MADPRLIDHPKLIBESIFRELRMAVMGANVLEAAVFPQEDADIPQIKNTNDPFAQGT 301
 QY 375 LTRQR-DMDXGLVLTSLIVLKSNTWMLDIVSTRMLGOYGFPLARVSGICIEDLCISYDC 433
 Db 302 KLSNHEIEKENG--LTSLAGKDPFLSITLFRKMSDEIGFIKMSIFPKHE--ISIEH 356
 QY 434 VATSEVSVSLSDSKTSRELIQASLHDHVELEK---IAIVLLQORAIISLIGNV 490
 Db 357 IESGIDNIGV-----IYSAEAL--ADRLFITREIKETLGEIEIVIEDLALISVYGGP 408
 QY 491 EGSLLIF-KIGRVLRKSGVNVQMSOGASKVMSLIVHSDAKALVEALHQAFFED 547
 Db 409 HKEILGSGKVLSTLNKLEIRTSLSQCAQELNLTIGVNNQIVETVVKGIYEGMNTD 466
 RESULT 11
 US-08-380-182-19
 / Sequence 19, Application US/08380182
 / Patent No. 5858749
 / GENERAL INFORMATION:
 / APPLICANT: Matthews, Benjamin F.
 / TITLE OF INVENTION: A Bifunctional Protein From Carrots
 / TITLE OF INVENTION: (Daucus carota) with Aspartokinase and Homoserine
 / TITLE OF INVENTION: Dehydratogenase Activities
 / NUMBER OF SEQUENCES: 23
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Janelle S. Graeter
 / STREET: Room 411, Bldg. 005, BARC-W
 / CITY: Beltsville
 / STATE: Maryland
 / COUNTRY: USA
 / ZIP: 20705
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: Patentin Release #1.0, Version #1.25
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/380,182
 / FILING DATE:
 / CLASSIFICATION: 435
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Graeter, Janelle S.
 / REGISTRATION NUMBER: 35,024
 / REFERENCE/DOCKET NUMBER: 0226.94
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 301-504-6629
 / TELEFAX: 301-504-5060
 / INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:
 LENGTH: 863 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-380-182-19

Query Match 14.8%; Score 416.5; DB 2; length 863;
 Best Local Similarity 26.3%; Pred. No. 9, 2e-33;
 Matches 126; Conservative 108; Mismatches 210; Indels 35; Gaps 16;

QY 86 VMKFGSSVSSAARMAEVALITPEEPFVVLSAMGKTN--NLLAEKAVGCCVI 142
 DB 36 IHKFGTGVSSSRIRHVAEIVVEDSEKRLVVASMSKVTDMVLLYKQSRDYES 95
 QY 143 HASEIEEMNVKSLHKTVDLG-LEXICNTSLVELEQLKGIAMKSLTPTSDYVSF 201
 DB 96 ALDAMEKHKLTFDLDDGLARFLRLQHDVNNKAMLRALYTAGHATSEFSDVVG 155
 QY 202 GECMSTRIFSAYLKIRYARQYDAFDIGITTFDFGNADLEATYPAVAKRLHGDW 261
 DB 156 GELMSAQLLSFVIRKGGDCNMWDRDY--LVNPAQSNQV-DPDYLESEKRLK-KWSS 211
 QY 262 PAIP--VTGFLGKMSKSAVITLGGSGDLATITIGALREIQWKVDVDTCPN 319
 DB 212 NQCQITVATGFIA-STPCNIPTTLKRKDSDFSAALMGALLRAGQVITWTVNGYSADPR 270
 QY 320 IYPAKTVPLTFEEATELAYFGAQLHPQSMRPAEGDIPVRKNSYNPAKPGILLTRQ 379
 DB 271 KVEEAVVVKTLSSQEMMSYFGANVLRHPTIIPVWRVDPVIVININLSAPGTMICRE 330
 QY 380 R--DMDXGLV--LTSIVLKSNTMLDIVSTRMGQYGFLLARVSGICIEDLCISVDCV 434
 DB 331 SVGETEDGKLSESHVGFATIDMLALINEGTGMAGVPGTASAIIGA--VKDVGANVMI 388
 QY 435 --ATSEVSVSILDPKISRELIIQASE-----LDHVELEKIAIVRLQPARISL 486
 DB 389 SQASSEHSICFAYPSESEV--KAVAKLARPRLQD--ANLSQVALI---PNCISLAT 440
 QY 487 IGNEQSSILLEKT-GRVLKSGVNVQMTSQGASKYMSLIYHSDAKALVEALHQAQF 544
 DB 441 VQGMKASTPGVSAITLFLNALAKANINRAIAQGCTEYINITVLSREDCVARALAVHSREY 499

RESULT 12
 US-09-134-001C-4160
 / Sequence 4160, Application US/09134001C
 / Patent No. 6380370
 / GENERAL INFORMATION:
 / APPLICANT: Lynn Doucette-Stamm et al
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 / TITLE OF INVENTION: EPIDERMIDS FOR DIAGNOSTICS AND THERAPEUTICS
 / FILE REFERENCE: GTC-007
 / CURRENT APPLICATION NUMBER: US/09/134,001C
 / CURRENT FILING DATE: 1998-08-13
 / PRIOR APPLICATION NUMBER: US 60/064,964
 / PRIOR FILING DATE: 1997-11-08
 / PRIOR APPLICATION NUMBER: US 60/055,779
 / PRIOR FILING DATE: 1997-08-14
 / NUMBER OF SEQ ID NOS: 5674
 / SEQ ID NO 4160
 / LENGTH: 463
 / TYPE: PRT
 / ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-4160

Query Match 14.8%; Score 415.5; DB 4; length 463;
 Best Local Similarity 27.0%; Pred. No. 4, 2e-33;
 Matches 127; Conservative 108; Mismatches 205; Indels 31; Gaps 13;
 QY 80 GDQLSVMKFGSSVSSAARMAEVALITPEEPFVVLSAMG-----KTNMLLA 132
 DB 2 GRKIMYAKFGSSVSTABQIKKYLTVNEDP-ERKIIIVSAPGRKHNDDIKTDLRL 60

QY 133 GERAVCGVIVHSEIEEMNVKSLHKTVDLGLEXICNTSLVELEQLKGIAMKSLT 191
 DB 61 YEKVLN-KLVNSKQCE---IIQRYADIVELG---IGNDLILINPTLEBYIKHLSKP 113
 QY 192 PRTSDYVSFGECSTRIFSAYLKIRYARQYDAFDIGITTFDFGNADLEATYPAVA 251
 DB 114 NRYLALISCGENFNAQLAQNNSOGIPYVISPKEAGLTVDLPOAAILDSAVNEYI 173
 QY 252 KALHGDWIDPAIPVVTGFLGKMSKSAVITLGGSGDLATITIGALREIQWKVDV 311
 DB 174 KL--RDYDEKLIIP--GPGVS-KQNYIVTFPPGSGDITGAIARGVASLYNFTDVS 227
 QY 312 GVLCDPNITPAKTVPLTFEEATELAYFGAQLHPQSMRPAEGDIPVRKNSYNPAK 371
 DB 228 GIYKANPNINNPDLIEITVREMERELSYAGFGVHDEALQPLKCRIPVVIKNTNRPND 287
 QY 372 PGILLTRQDMXGLVITLSIVLKSNTMLDIVSTRMGQYGFLLARVSGICIEDLCISV 431
 DB 288 KGTYLHDEKIDSKNVI-SGISCDKGFVVIKYLKMLRLVGFTRKILGV--LEEFNISF 344
 QY 432 DCVATSEVSVSILDPKISRELIIQASELDHVELEKIAIVRLQPARISLIG-NV 490
 DB 345 DHMPSGIDNISITMETNQIGKE-----SQVLNMRIRCEVDELSIDHDLAVIMTVEGM 399
 QY 491 EQSSLILEKTGRVLKSGVNVQMTSQGASKYMSLIYHSDAKALVEALHQAQF 541
 DB 400 NVVGTVASKITHALSESINILININQASSETSMWFGIHEADAKAVLSTYE 450

RESULT 13
 US-08-532-828B-3
 / Sequence 3, Application US/08532828B
 / Patent No. 5686671
 / GENERAL INFORMATION:
 / APPLICANT: SUGIMOTO, Masakazu
 / APPLICANT: OGAWA, Yuri
 / APPLICANT: SUZUKI, Tomoko
 / APPLICANT: TANAKA, Akiko
 / APPLICANT: YATSUI, Hiroshi
 / TITLE OF INVENTION: MUTANT ASPARTOKINASE GENE
 / NUMBER OF SEQUENCES: 24
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
 / STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
 / CITY: ARLINGTON
 / STATE: VA
 / COUNTRY: USA
 / ZIP: 22202
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: MS-DOS TEXT EDITOR
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/532,828B
 / FILING DATE: 27-OCT-1995
 / CLASSIFICATION: 435
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: JP 5-101450
 / FILING DATE: 27-APR-1993
 / ATTORNEY/AGENT INFORMATION:
 / NAME: NORMAN F. OBLON
 / REGISTRATION NUMBER: 24,618
 / REFERENCE/DOCKET NUMBER: 10-764-0 PCT
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 703-413-3000
 / TELEFAX: 703-413-2220
 / INFORMATION FOR SEQ ID NO: 3:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 421 amino acids
 / TYPE: amino acid
 / TOPOLOGY: linear

MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Corynebacterium glutamicum
STRAIN: ATCC13639
US-08-532-828B-3

Query Match
Best Local Similarity 28.5%; Score 414.5; DB 1; Length 421;
Matches 135; Conservative 83; Mismatches 169; Indels 87; Gaps 14;

14.7%; Score 414.5; DB 1; Length 421;
Best Local Similarity 28.5%; Score 414.5; DB 1; Length 421;
Matches 135; Conservative 83; Mismatches 169; Indels 87; Gaps 14;

85 VYMKFGSSVSSAARMAEVALITLTPPE--ERPVVVLSAMGKTTNNLLAGEKAVGCVI 142
4 VVQKYGSSLSAEIRVNAERIVATKAGNDVYVCSAGDITDEL-----51
143 HVSEIEMNMVXSLHIKTVDLGLPXICNTSLYELEQLKGIAMMKELTPRTSDYVSG 202
52 -----ELAAVNPVPAREM-----DMLLTAG 73
203 ECMSTRIFSAVINKIRVAKQYDAPDIGITTFDEFGNADILEATYPAVAKLHGMIDP 262
74 ERISNALVMAIESIGAEKQSFQSGAVLTTERHGNARIVDTTPGRVREAL-----DEG 128
263 AIPVVTGFLGKMGSGAVTTLGRGSDLTATTIGKALREIQWKVDVGLTCDPNIYP 322
129 KICIVAGFGQVKKETRDVTLGRGSDTTAVALAALNADVCEIYSDVDGYTADPRIVP 188
323 HAKTVPYLTFEEATELAFGAOVLHPQSMRAREBDIPVRKNSINPAKGTILTT-RQRD 361
189 NQKLEKLSFEEMLELAAGSKILVRSVEYARAFNPFLKRSYS-NDPGLIAGSMED 247
382 MDXGLVLTSTIVL--KSNVTMLDIYSTRLMGQYGLARVSGICYIEDLCISVDCVATSE 438
248 IVEEAVLTGVAITDSEAKVTYVLGISD-----KPGEMAKV--FRALDAEINIDMTVLQNV 300
439 VSVSVSLDPSKIMSELIQOASELD--HYVELEKIAL-----VRLIQRAIISLIG-N 489
301 SSVEDG-----TTDITFTCPRADGRRAHEILKLDVQGWNTVLYDDQVKXSLVAG 353
490 VEGSSLIIEKTRVLRKSGVNVOMISQASKYNNMLIYHSDPAKALVEALHQAIF 543
354 MKSHPGVTAPEFMAELADVNVTIELIS--TSIRIRSVLIREDDLDAAARALHBOF 405

RESULT 14
US-08-532-828B-4
Sequence 4, Application US/08532828B
Patent No. 5688671
GENERAL INFORMATION:
APPLICANT: SUGIMOTO, Masakazu
APPLICANT: OGAWA, YURI
APPLICANT: SUZUKI, Tomoko
APPLICANT: TANAKA, AKIKO
APPLICANT: MATSUI, Hiroshi
TITLE OF INVENTION: MUTANT ASPARTOKINASE GENE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-DOS TEXT EDITOR
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532,828B
FILING DATE: 27-OCT-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-101450
FILING DATE: 27-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: NORMAN F. OBLON
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-764-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 421 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Corynebacterium glutamicum
STRAIN: AJ3463
US-08-532-828B-4

Query Match
Best Local Similarity 28.5%; Score 414.5; DB 1; Length 421;
Matches 135; Conservative 83; Mismatches 169; Indels 87; Gaps 14;

14.7%; Score 414.5; DB 1; Length 421;
Best Local Similarity 28.5%; Score 414.5; DB 1; Length 421;
Matches 135; Conservative 83; Mismatches 169; Indels 87; Gaps 14;

85 VYMKFGSSVSSAARMAEVALITLTPPE--ERPVVVLSAMGKTTNNLLAGEKAVGCVI 142
4 VVQKYGSSLSAEIRVNAERIVATKAGNDVYVCSAGDITDEL-----51
143 HVSEIEMNMVXSLHIKTVDLGLPXICNTSLYELEQLKGIAMMKELTPRTSDYVSG 202
52 -----ELAAVNPVPAREM-----DMLLTAG 73
203 ECMSTRIFSAVINKIRVAKQYDAPDIGITTFDEFGNADILEATYPAVAKLHGMIDP 262
74 ERISNALVMAIESIGAEKQSFQSGAVLTTERHGNARIVDTTPGRVREAL-----DEG 128
263 AIPVVTGFLGKMGSGAVTTLGRGSDLTATTIGKALREIQWKVDVGLTCDPNIYP 322
129 KICIVAGFGQVKKETRDVTLGRGSDTTAVALAALNADVCEIYSDVDGYTADPRIVP 188
323 HAKTVPYLTFEEATELAFGAOVLHPQSMRAREBDIPVRKNSINPAKGTILTT-RQRD 361
189 NQKLEKLSFEEMLELAAGSKILVRSVEYARAFNPFLKRSYS-NDPGLIAGSMED 247
382 MDXGLVLTSTIVL--KSNVTMLDIYSTRLMGQYGLARVSGICYIEDLCISVDCVATSE 438
248 IVEEAVLTGVAITDSEAKVTYVLGISD-----KPGEMAKV--FRALDAEINIDMTVLQNV 300
439 VSVSVSLDPSKIMSELIQOASELD--HYVELEKIAL-----VRLIQRAIISLIG-N 489
301 SSVEDG-----TTDITFTCPRADGRRAHEILKLDVQGWNTVLYDDQVKXSLVAG 353
490 VEGSSLIIEKTRVLRKSGVNVOMISQASKYNNMLIYHSDPAKALVEALHQAIF 543
354 MKSHPGVTAPEFMAELADVNVTIELIS--TSIRIRSVLIREDDLDAAARALHBOF 405

RESULT 15
US-08-700-359-9
Sequence 9, Application US/08700359
Patent No. 5766925
GENERAL INFORMATION:
APPLICANT: SUGIMOTO, MASAKAZU
APPLICANT: USUDA, YOSHITARO
APPLICANT: SUZUKI, TOMOKO
APPLICANT: TANAKA, AKIKO
APPLICANT: MATSUI, HIROSHI
TITLE OF INVENTION: METHOD OF PRODUCING L-LYSINE
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-DOS TEXT EDITOR
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/700,359
FILING DATE: 08-OCT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-35019
FILING DATE: 04-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: NORMAN F. OBLON
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-819-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-2220
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 421
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-700-359-9

Query Match 14.7%; Score 414.5; DB 1; Length 421;
Best Local Similarity 28.5%; Pred. No. 4.5e-33;
Matches 135; Conservative 93; Mismatches 169; Indels 97; Gaps 14;

QY 85 VMKFGSSVSSAARMAEAGLITPE--ERYVVISAMGKTTNNLLAGEKAVGCGVI 142
DB 4 VVCKYGGSSLESARIRNVAERIVATKAGNDVYVCSAMGDTTDEL----- 51
QY 143 HVSEIEMNMVXSLHITVDLGLPXICNTSLYEQLKIMMKELTPRTSDYLSFG 202
DB 52 -----ELAAVNPVPAKEM-----DMLLTNG 73
QY 203 ECMSTRIFSAVINKIRYKARQYDAFDIGITTDGFGNADILEATYPAVAKRLHGDWIODP 262
DB 74 ERISNALVMAIESLGAESFTGSCAGVLTTERHGNARIVDVTPGRVREAL----DEG 128
QY 263 AIPVVTGFLGKMGSGAVTLTGRGSDLTATTIGKALGLREIQVWKVDVGLTCDPNIYE 322
DB 129 KICIVAGPQGVNKTGRTVTLTGRGSDTLTVALAALNADVCEIYSDVDGVYADPRIVE 188
QY 323 HAKTVPLYTEFEATELAVFGAQLHPQMRPAREGDIPIYVKNKSNPKAPGTLIT--RORD 381
DB 189 NAGLEKLSFEEMELIAVSGKILVRSVEYARAPVPLKVRSSYS--NDPGTLIAGSMED 247
QY 382 MDKGLVLTSLVL--KSNVTMLDIVSTRMLGOYGFARVSGICIEDLCISVDCVATSE 438
DB 248 IPVEEAVLTGATDKSARATVIGISD---KPEGAKV--PRALDAEININDVYLQNV 300
QY 439 VSVSVSLDPSKIMRELIQASELD--HVEEELKIAL-----VRLLOPRAIISLIG-N 489
DB 301 SSVEDG-----TTDTITFCPRADGRAMEILKLGQGVNWTNVLVDQGVKSLVGAG 353
QY 490 VEOGSLIEKTVLRKSGVNVQVMTISQASKVNSLIVHSDAKALVEALHQAQF 543
DB 354 MKSHPGVTAEFMEALRDVNVNIELIS--TSEIRISVLIREDDDAAARALHROF 405

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 13:17:42 ; Search time 67 Seconds
[without alignments]
2164.403 Million cell updates/sec

Title: US-09-890-813-6
Perfect score: 2813
Sequence: 1 MAIPVSAAPRLVPSIRP.....QAFEDVLQVEANILLVG 560

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 25895539 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	2809	99.9	560	9 US-09-890-813-6	Sequence 6, Appli
2	2678.5	95.2	555	9 US-09-890-813-8	Sequence 8, Appli
3	1981.5	70.4	415	12 US-10-425-114-50163	Sequence 50163, A
4	1839	65.4	564	12 US-10-424-559-170907	Sequence 170907, A
5	1835	65.2	564	9 US-09-890-813-17	Sequence 17, Appli
6	1826.5	64.9	493	12 US-10-425-114-57733	Sequence 52733, A
7	1754.5	62.4	561	12 US-10-424-559-170902	Sequence 170902, A
8	1623.5	57.9	439	9 US-09-890-813-16	Sequence 16, Appli
9	1587.5	56.4	426	12 US-10-425-114-70672	Sequence 70672, A
10	1573	55.9	426	12 US-10-425-114-39779	Sequence 39779, A
11	1222.5	43.5	326	12 US-10-425-114-39716	Sequence 39716, A
12	992.5	35.3	249	12 US-10-425-114-43083	Sequence 43083, A
13	972	34.6	221	12 US-10-425-114-54453	Sequence 54453, A
14	890.5	31.7	281	9 US-09-890-813-12	Sequence 12, Appli
15	762.5	27.1	226	12 US-10-425-114-46004	Sequence 46004, A

16	671.5	23.9	464	15 US-10-369-493-11215	Sequence 11215, A
17	666	23.7	152	9 US-09-890-813-4	Sequence 4, Appli
18	662	23.5	473	15 US-10-369-493-21474	Sequence 21474, A
19	642.5	22.8	462	15 US-10-369-493-978	Sequence 978, App
20	607	21.6	473	15 US-10-369-493-584	Sequence 584, App
21	604.5	21.5	465	15 US-10-369-493-8972	Sequence 8972, App
22	586.5	20.8	178	12 US-10-425-114-51028	Sequence 51028, A
23	578.5	20.6	817	15 US-10-369-493-10372	Sequence 10372, A
24	575	20.4	447	15 US-10-369-493-21212	Sequence 21212, A
25	572	20.3	449	15 US-09-899-339-23	Sequence 23, Appli
26	569	20.2	449	15 US-10-369-493-23654	Sequence 23654, A
27	552	19.6	448	15 US-10-369-493-503	Sequence 503, App
28	535	19.0	481	15 US-10-369-493-10467	Sequence 10467, A
29	508	18.1	811	12 US-10-282-122A-48795	Sequence 48795, A
30	495.5	17.6	489	15 US-10-369-493-22563	Sequence 22563, A
31	491.5	17.5	527	15 US-10-369-493-1741	Sequence 1741, Ap
32	490.5	17.4	815	12 US-10-282-122A-66881	Sequence 66881, A
33	483.5	17.2	825	12 US-10-282-122A-77839	Sequence 77839, A
34	478.5	17.0	580	12 US-10-282-122A-56011	Sequence 56011, A
35	477.5	17.0	820	12 US-10-282-122A-42805	Sequence 42805, A
36	475.5	16.9	820	15 US-10-369-493-686	Sequence 686, App
37	475.5	16.9	815	12 US-10-282-122A-58074	Sequence 58074, A
38	470.5	16.7	822	15 US-10-282-122A-59647	Sequence 59647, A
39	469.5	16.7	820	12 US-10-282-122A-74838	Sequence 74838, A
40	469.5	16.7	820	12 US-10-282-122A-74838	Sequence 74838, A
41	469.5	16.7	820	12 US-10-282-122A-75391	Sequence 75391, A
42	468.5	16.7	454	12 US-10-282-122A-73501	Sequence 73501, A
43	466	16.6	818	15 US-10-369-493-19453	Sequence 19453, A
44	465.5	16.5	812	15 US-10-369-493-15660	Sequence 15660, A
45	465.5	16.5	812	15 US-10-369-493-16051	Sequence 16051, A

ALIGNMENTS

RESULT 1
US-09-890-813-6
Sequence 6, Application US/09890813
Publication No. US20020183486A1
GENERAL INFORMATION:
APPLICANT: E. I. du Pont de Nemours and Company
TITLE OF INVENTION: Aspartate Kinase
FILE REFERENCE: BB1430 PCT
CURRENT APPLICATION NUMBER: US/09/890,813
PRIOR APPLICATION NUMBER: 60/172294
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Microsoft Office 97
SEQ ID NO 6
LENGTH: 560
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: UNSURE
LOCATION: (168)
NAME/KEY: UNSURE
LOCATION: (384)
US-09-890-813-6
Query Match 99.9%; Score 2809, DB 9, Length 560;
Best Local Similarity 100.0%; Pred. No. 1.4e-267;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 MAIPVSAAPRLVPSIRPSSGVRGLACGTGTGGRGAGLSMVVADTSRRKAD 60
Db 1 MAIPVSAAPRLVPSIRPSSGVRGLACGTGTGGRGAGLSMVVADTSRRKAD 60
Cy GGGGVGAPVLGGLGEGIGDOLSVYMKFGSSVSAPRABVAGLIITPEEPVYVL 120
Db GGGGVGAPVLGGLGEGIGDOLSVYMKFGSSVSAPRABVAGLIITPEEPVYVL 120
Cy 121 AMGKTNNLLGKAKVGVHVEIEBMNVKSLHKTYDELGLPXCITSLEYELQL 180

Db 121 AMGKTTNNLLAGEKAVGCGVIVHSEIEMNNVKSLLIKITVDELGLPXCINTSLYELEQI 180
 Qy 181 LKGIAMMKELTPRTSDYLVSFGECMSTRIFSAVYANKIRVAKROYDADFIGITTFDEFGNA 240
 Db 181 LKGIAMMKELTPRTSDYLVSFGECMSTRIFSAVYANKIRVAKROYDADFIGITTFDEFGNA 240
 Qy 241 DILEATYPAVARLHGDWIODPAIPVVTGFLGKMGKSAVTTLGRGSDLTATTIGKALG 300
 Db 241 DILEATYPAVARLHGDWIODPAIPVVTGFLGKMGKSAVTTLGRGSDLTATTIGKALG 300
 Qy 301 LREIOWKDVGVLTCDPNIPYHAKTVYLLFEEATELAYFGAQLHPQSMRPARBEDI 360
 Db 301 LREIOWKDVGVLTCDPNIPYHAKTVYLLFEEATELAYFGAQLHPQSMRPARBEDI 360
 Qy 361 VPKVSYNPKARPGTLITRQDMXGLVLTSLVLSKNTMLDIVSTMLGQYGLARVSG 420
 Db 361 VPKVSYNPKARPGTLITRQDMXGLVLTSLVLSKNTMLDIVSTMLGQYGLARVSG 420
 Qy 421 ICYIEDLCISVDCVATSEVSVSLDPKIMSRRELIOQASELDHVEELEKIALVRLIQ 480
 Db 421 ICYIEDLCISVDCVATSEVSVSLDPKIMSRRELIOQASELDHVEELEKIALVRLIQ 480
 Qy 481 RAIIISLIGNVEQSSLLIEKTGRVLRKSGVNVQMISQASKNMSLIYHDSAKALVEALH 540
 Db 481 RAIIISLIGNVEQSSLLIEKTGRVLRKSGVNVQMISQASKNMSLIYHDSAKALVEALH 540
 Qy 541 QAFEDDVLQVYAEENLVG 560
 Db 541 QAFEDDVLQVYAEENLVG 560

RESULT 2
 US-09-890-813-6
 / Sequence 8, Application US/09890813
 / Publication No. US20020183486A1
 / GENERAL INFORMATION:
 / APPLICANT: E. I. du Pont de Nemours and Company
 / TITLE OF INVENTION: Aspartate Kinase
 / FILE REFERENCE: BA1430 PCT
 / CURRENT APPLICATION NUMBER: US/09/890, 813
 / CURRENT FILING DATE: 2001-08-02
 / PRIOR APPLICATION NUMBER: 60/172944
 / PRIOR FILING DATE: 1999-12-21
 / NUMBER OF SEQ ID NOS: 24
 / SOFTWARE: Microsoft Office 97
 / SEQ ID NO 8
 / LENGTH: 565
 / TYPE: PRT
 / ORGANISM: Zea mays
 / US-09-890-813-6

Query Match 95.2%; Score 2678.5; DB 9; Length 555;
 Best Local Similarity 96.8%; Pred. No. 1e-254;
 Matches 543; Conservative 1; Mismatches 10; Indels 7; Gaps 4;

Qy 1 MAIPVSAAPRRLLVPSIPASSGAVRGLACGTITGPRGAGLSMVAADSTSRKAD 60
 Db 1 MAIPVSAAPRRLLVPSIPASSGAVRGLACGTITGPRGAGLSMVAADSTSRKAD 60
 Qy 61 GGGGVGAPVGLGLMGEG:GDOLSVYMKFGSSVSAARMAEVALITTFEERPVVLS 120
 Db 61 GGGGVGAPVGLGLMGEG:GDOLSVYMKFGSSVSAARMAEVALITTFEERPVVLS 120
 Qy 121 AMGKTTNNLLAGEKAVGCGVIVHSEIEMNNVKSLLIKITVDELGLP-XICNTSLYELEQ 179
 Db 121 AMGKTTNNLLAGEKAVGCGVIVHSEIEMNNVKSLLIKITVDELGLP-XICNTSLYELEQ 180
 Qy 180 LKGIAMMKELTPRTSDYLVSFGECMSTRIFSAVYANKIRVAKROYDADFIGITTFDEFGN 239
 Db 181 LKGIAMMKELTPRTSDYLVSFGECMSTRIFSAVYANKIRVAKROYDADFIGITTFDEFGN 240
 Qy 240 ADILEATYPAVARLHGDWIODPAIPVVTGFLGKMGKSAVTTLGRGSDLTATTIGKAL 299

Db 241 ADILEATYPAVARLHGDWIODPAIPVVTGFLGKMGKSAVTTLGRGSDLTATTIGKAL 300
 Qy 300 GLREIOWKDVGVLTCDPNIPYHAKTVYLLFEEATELAYFGAQLHPQSMRPARBEDI 359
 Db 301 GLREIOWKDVGVLTCDPNIPYHAKTVYLLFEEATELAYFGAQLHPQSMRPARBEDI 360
 Qy 360 PVKVSYNPKARPGTLITRQDMXGLVLTSLVLSKNTMLDIVSTMLGQYGLARV 419
 Db 361 PVKVSYNPKARPGTLITRQDMXGLVLTSLVLSKNTMLDIVSTMLGQYGLARV 418
 Qy 420 GICIEDLCISVDCVATSEVSVSLDPKIMSRRELIOQASELDHVEELEKIALVRLIQ 479
 Db 419 AI--FEDLCISVDCVATSEVSVSLDPKIMSRRELIOQ--ELDHVEELEKIALVRLIQ 474
 Qy 480 QRAIIISLIGNVEQSSLLIEKTGRVLRKSGVNVQMISQASKNMSLIYHDSAKALVEAL 539
 Db 475 QRAIIISLIGNVEQSSLLIEKTGRVLRKSGVNVQMISQASKNMSLIYHDSAKALVEAL 534
 Qy 540 HQAFEDDVLQVYAEENLVG 560
 Db 535 HQAFEDDVLQVYAEENLVG 555

RESULT 3
 US-10-425-114-50163
 / Sequence 50163, Application US/10425114
 / Publication No. US20040034888A1
 / GENERAL INFORMATION:
 / APPLICANT: Liu, Jingdong
 / APPLICANT: Zhou, Yihua
 / APPLICANT: Kowalic, David K.
 / APPLICANT: Screen, Steven E
 / APPLICANT: Tabaska, Jack E
 / APPLICANT: Cao, Yongwei
 / TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 / FILE REFERENCE: 38-21(5313)B
 / CURRENT APPLICATION NUMBER: US/10/425, 114
 / CURRENT FILING DATE: 2003-04-28
 / NUMBER OF SEQ ID NOS: 73128
 / SEQ ID NO 50163
 / LENGTH: 415
 / TYPE: PRT
 / ORGANISM: Zea mays
 / FEATURE:
 / OTHER INFORMATION: Clone ID: 700553382_F11.pep
 / US-10-425-114-50163

Query Match 70.4%; Score 1981.5; DB 12; Length 415;
 Best Local Similarity 95.7%; Pred. No. 3.4e-186;
 Matches 403; Conservative 1; Mismatches 10; Indels 7; Gaps 4;

Qy 141 VHVSEIEMNNVKSLLIKITVDELGLP-XICNTSLYELEQLKGIAMMKELTPRTSDYLV 139
 Db 1 VHVSEIEMNNVKSLLIKITVDELGLPXSIVQDMLELEQLKGIAMMKELTPRTSDYLV 60
 Qy 200 SFGECMSTRIFSAVYANKIRVAKROYDADFIGITTFDEFGNADILEATYPAVARLHGDW 259
 Db 61 SFGECMSTRIFSAVYANKIRVAKROYDADFIGITTFDEFGNADILEATYPAVARLHGDW 120
 Qy 260 QDPALPVVTGFLGKMGKSAVTTLGRGSDLTATTIGKALGLREIOWKDVGVLTCDPN 319
 Db 121 QDPALPVVTGFLGKMGKSAVTTLGRGSDLTATTIGKALGLREIOWKDVGVLTCDPN 180
 Qy 320 IYPAKTVYLLFEEATELAYFGAQLHPQSMRPARBEDI PVKVSYNPKARPGTLITRQ 379
 Db 181 IYPAKTVYLLFEEATELAYFGAQLHPQSMRPARBEDI PVKVSYNPKARPGTLITRQ 240
 Qy 380 RMDXGLVLTSLVLSKNTMLDIVSTMLGQYGLARVSGICIEDLCISVDCVATSEV 439
 Db 241 RMDXGLVLTSLVLSKNTMLDIVSTMLGQYGLARVSGICIEDLCISVDCVATSEV 296

```

QY 440 SVSVSLDPSKIMRELIIQASSELDRVVEELEKAIYRLIQORAILIIGNVEQSSLIIEK 499
DB 297 SVSVSLDPSKIMRELIIQ--ELDHVVEELEKAIYRLIQORAILIIGNVEQSSLIIEK 354
QY 500 TGSVLKRSQVNVQMISQASKVMMSLIYHDSDAKALVEALHOAFEDDVLISQVEANLILV 559
DB 355 TGSVLKRSQVNVQMISQASKVMMSLIYHDSDAKALVEALHOAFEDDVLISQVEANLILV 414
QY 560 G 560
DB 415 G 415

```

RESULT 4

```

US-10-424-599-170907
; Sequence 170907, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 170907
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(564)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_125343C.1.Pep
US-10-424-599-170907

```

```

Query Match 65.4%; Score 1839; DB 12; Length 564;
Best Local Similarity 71.0%; Pred. No. 6,3e-172;
Matches 391; Conservative 40; Mismatches 80; Indels 40; Gaps 7;

```

```

QY 37 GPRGARGLSMVVADSNRRRAKQADGDVLAGPVL----- 71
DB 10 GVQGLKLAIVMSVRSLSLHCKSQI--GFALGAPVCARRWGNRVAFSTTCKASTDYIE 67
QY 72 -----GGLMGEGLDQSLVVMKFGGSSVSSAARMAVAGLITPPEERPVVLSAMGKT 125
DB 68 KNATENGVMSSSE-ETSFCTVMKFGSSVSAADRMKEVATLILSPEERPIVVLAMGKT 126
QY 126 TNNLLAGKRAVAGCGVIHSEIEEMNMVSLHIKIVDELGLP-XICNTSLVELQLLKI 184
DB 127 TNNLLAGKRAVAGCGVINVSIEELCFIKDLHRTVDQGVDSVISKHLIELQLLKI 186
QY 185 AMKELTPRTSDYLVSFGECMSTRIFSAYLNKIRVAKQYDAFDIGFITTDFGNADILE 244
DB 187 AMKELTPRTSDYLVSFGECMSTRIFSAYLNKIRVAKQYDAFDIGFITTDFGNADILE 246
QY 245 ATYPAVAKRLHGMIDDPALPVYTGFLGKMGSGAVTTLGRGSDLTATTTKALGLREI 304
DB 247 ATYPAVAKRLHGMIDDPALPVYTGFLGKMGSGAVTTLGRGSDLTATTTKALGLREI 306
QY 305 QVMKDVGVLTCDPNITPRAKTPVYLTFEATLAFGAQVLAHPOSMRPAARECDIPVRVK 364
DB 307 QVMKDVGVLTCDPNITPRAKTPVYLTFEATLAFGAQVLAHPOSMRPAARECDIPVRVK 366
QY 365 NSYNPRAFGTLITRORDMXGLVLTSLVLSKNTVMDIVSTRMLGQYGFILARVSGICYI 424
DB 367 NSYNPRAFGTLITRORDMXGLVLTSLVLSKNTVMDIVSTRMLGQYGFILARVSGICYI 422
QY 425 EDLCISVDCAVTSVSVSLDPSKIMRELIIQASSELDRVVEELEKAIYRLIQORAILI 484

```

```

DB 423 EELGISVDVATSEVSVSLTIDPSKIMRELIIQASSELDRVVEELEKAIYRLIQORAILI 482
QY 485 SLIGNVEQSSLIIEKTRGVLRKSGVNVQMISQASKVMMSLIYHDSDAKALVEALHOAF 544
DB 483 SLIGNVEQSSLIIEKTRGVLRKSGVNVQMISQASKVMMSLIYHDSDAKALVEALHOAF 542
QY 545 EDDVLISQVEAE 555
DB 543 ESE-ISELEMD 552

```

RESULT 5

```

US-09-890-813-17
; Sequence 17, Application US/09890813
; Publication No. US20020183486A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Aspartate Kinase
; FILE REFERENCE: BB1430 PCT
; CURRENT APPLICATION NUMBER: US/09/890,813
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/172944
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Glycine max
US-09-890-813-17

```

```

Query Match 65.2%; Score 1835; DB 9; Length 564;
Best Local Similarity 70.6%; Pred. No. 1.6e-171;
Matches 389; Conservative 42; Mismatches 80; Indels 40; Gaps 7;

```

```

QY 37 GPRGARGLSMVVADSNRRRAKQADGDVLAGPVL----- 71
DB 10 GVQGLKLAIVMSVRSLSLHCKSQI--GFALGAPVCARRWGNRVAFSTTCKASTDYIE 67
QY 72 -----GGLMGEGLDQSLVVMKFGGSSVSSAARMAVAGLITPPEERPVVLSAMGKT 125
DB 68 KNATENGVMSSSE-ETSFCTVMKFGSSVSAADRMKEVATLILSPEERPIVVLAMGKT 126
QY 126 TNNLLAGKRAVAGCGVIHSEIEEMNMVSLHIKIVDELGLP-XICNTSLVELQLLKI 184
DB 127 TNNLLAGKRAVAGCGVINVSIEELCFIKDLHRTVDQGVDSVISKHLIELQLLKI 186
QY 185 AMKELTPRTSDYLVSFGECMSTRIFSAYLNKIRVAKQYDAFDIGFITTDFGNADILE 244
DB 187 AMKELTPRTSDYLVSFGECMSTRIFSAYLNKIRVAKQYDAFDIGFITTDFGNADILE 246
QY 245 ATYPAVAKRLHGMIDDPALPVYTGFLGKMGSGAVTTLGRGSDLTATTTKALGLREI 304
DB 247 ATYPAVAKRLHGMIDDPALPVYTGFLGKMGSGAVTTLGRGSDLTATTTKALGLREI 306
QY 305 QVMKDVGVLTCDPNITPRAKTPVYLTFEATLAFGAQVLAHPOSMRPAARECDIPVRVK 364
DB 307 QVMKDVGVLTCDPNITPRAKTPVYLTFEATLAFGAQVLAHPOSMRPAARECDIPVRVK 366
QY 365 NSYNPRAFGTLITRORDMXGLVLTSLVLSKNTVMDIVSTRMLGQYGFILARVSGICYI 424
DB 367 NSYNPRAFGTLITRORDMXGLVLTSLVLSKNTVMDIVSTRMLGQYGFILARVSGICYI 422
QY 425 EDLCISVDCAVTSVSVSLDPSKIMRELIIQASSELDRVVEELEKAIYRLIQORAILI 484
DB 483 SLIGNVEQSSLIIEKTRGVLRKSGVNVQMISQASKVMMSLIYHDSDAKALVEALHOAF 544
QY 485 EDDVLISQVEAE 555

```

Db 543 ESE-ISELEMD 552

RESULT 6
US-10-425-114-52733
Sequence 52733, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 52733
LENGTH: 493
TYPE: PR
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: 700556108_Flt.pep
US-10-425-114-52733

Query Match 64.9%; Score 1826.5; DB 12; Length 493;
Best Local Similarity 77.7%; Pred. No. 8.7e-171;
Matches 377; Conservative 39; Mismatches 62; Indels 7; Gaps 5;

QY 72 GCGMGBGLDQLSVNKKFGSSSVSAAAEVAGLITPPEEPVYVLSANGKTTNNLL 131
DB 3 GWSVSSBG-ETSFVCKMFGSSSVASADRMKEVATLILSPPEEPVYVLSANGKTTNNLL 61
QY 132 AGEKAVGCGVIHYSEIEEMNMVKSILIKTVDELGLP-XICNTSLYEELQILKGIAMMEL 190
DB 62 AGEKAVGCGVIHYSEIEEMNMVKSILIKTVDELGLP-XICNTSLYEELQILKGIAMMEL 121
QY 191 TPRTSDYLVSPFGCMSTRIFSAYLKIRKARQYADPDIGFTTTFEFGNADILETYAV 250
DB 122 TKCTODYLVSPFGCMSTRIFSAYLKIRKARQYADPDIGFTTTFEFGNADILETYAV 161
QY 251 AKLHGDMLDPAIPVYTGFLGKMGSGAVTTLGRGSDLTATTTGKALGPEIQWMDV 310
DB 182 AKLHGDMLDPAIPVYTGFLGKMGSGAVTTLGRGSDLTATTTGKALGPEIQWMDV 241
QY 311 DGVLTCDPNITPAKTVPYLTPEATELAYFGAQLVHPOSMPAREGDIPIRVKNSYNPK 370
DB 242 DGVLTCDPNITPAKTVPYLTPEATELAYFGAQLVHPOSMPAREGDIPIRVKNSYNPK 301
QY 371 APGTLTTRORDMXXGLVLTSLYKSNVTMLDIVSTRMGQYGFARVSGLCYIEDLCIS 430
DB 302 APGTLTTRORDMXXGLVLTSLYKSNVTMLDIVSTRMGQYGFARVSGLCYIEDLCIS 357
QY 431 VDCVATSEVSVSLDPSKTIWRELLIOASELDHVEEIEKTAIVNLLQNRRIISLIINV 490
DB 358 VDCVATSEVSVSLDPSKTIWRELLIOASELDHVEEIEKTAIVNLLQNRRIISLIINV 417
QY 491 EGSGLIEKTRGVLRKSGVNVQMIISQASGVNMSLIVHSDAKALVEALHQAFFEDVLS 550
DB 418 EGSGLIEKTRGVLRKSGVNVQMIISQASGVNMSLIVHSDAKALVEALHQAFFEDVLS 476
QY 551 QVEAE 555
DB 477 ELEMND 481

RESULT 7
US-10-424-599-170902
Sequence 170902, Application US/10424599

Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 170902
LENGTH: 561
TYPE: PR
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_125339C.1.pep
US-10-424-599-170902

Query Match 62.4%; Score 1754.5; DB 12; Length 561;
Best Local Similarity 71.9%; Pred. No. 1.3e-163;
Matches 361; Conservative 52; Mismatches 74; Indels 15; Gaps 6;

QY 55 RAKQADGGDGVLAGAVIGLGMGBGLDQLSVNKKFGSSSVSAAAEVAGLITPPEEP 114
DB 62 KATESDVVEGSG-----GFATEMS YTCVMKFGSSVAMERKEVANTLILSPPEEP 114
QY 115 PIVYLSANGKTTNNLLAGEKAVGCGVIHYSEIEEMNMVKSILIKTVDELGLP-XICNTS 173
DB 115 PIVYLSANGKTTNNLLAGEKAVGCGVIHYSEIEEMNMVKSILIKTVDELGLP-XICNTS 174
QY 174 LYEELQILKGIAMMELTPRTSDYLVSPFGCMSTRIFSAYLKIRKARQYADPDIGFTT 233
DB 175 LYEELQILKGIAMMELTPRTSDYLVSPFGCMSTRIFSAYLKIRKARQYADPDIGFTT 234
QY 234 TFEFGNADILETYAVAKRLHGDMLDPAIPVYTGFLGKMGSGAVTTLGRGSDLTAT 293
DB 235 TFEFGNADILETYAVAKRLHGDMLDPAIPVYTGFLGKMGSGAVTTLGRGSDLTAT 294
QY 294 TIGKALGPEIQWMDVGVLTCDPNITPAKTVPYLTPEATELAYFGAQLVHPOSMP 353
DB 295 TIGKALGPEIQWMDVGVLTCDPNITPAKTVPYLTPEATELAYFGAQLVHPOSMP 354
QY 354 AREGDIPIRVKNSYNPKAPGTLITRORDMXXGLVLTSLYKSNVTMLDIVSTRMGQY 413
DB 355 AREGDIPIRVKNSYNPKAPGTLITRORDMXXGLVLTSLYKSNVTMLDIVSTRMGQY 412
QY 414 FLARVSGLCYIEDLCISVDCVATSEVSVSLDPSKTIWRELLIOASELDHVEEIEKIA 473
DB 413 FLARVSGLCYIEDLCISVDCVATSEVSVSLDPSKTIWRELLIOASELDHVEEIEKIA 468
QY 474 IYRLIQORAIISLIINVHSDAKALVEALHQAFFEDVLS 533
DB 469 IYRLIQORAIISLIINVHSDAKALVEALHQAFFEDVLS 528
QY 534 ACVEALHQAFFEDVLSQVEAE 555
DB 529 QCVRALHQAFFESF-ISELEND 549

RESULT 8
US-09-890-813-16
Sequence 16, Application US/09890813
Publication No. US20020183486A1
GENERAL INFORMATION:
APPLICANT: E.I. du Pont de Nemours and Company
TITLE OF INVENTION: Aspartate Kinase
FILE REFERENCE: B81430 PCT
CURRENT APPLICATION NUMBER: US/09/890,813
CURRENT FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: 60/172944
PRIOR FILING DATE: 1999-12-21

NUMBER OF SEQ ID NOS: 24
 SOFTWARE: Microsoft Office 97
 SEQ ID NO 16
 LENGTH: 439
 TYPE: PRT
 ORGANISM: *Triticum aestivum*
 US-09-890-813-16

Query Match 57.9%; Score 1629.5; DB 9; Length 439;
 Best Local Similarity 77.7%; Pred. No. 1.9e-151;
 Matches 331; Conservative 37; Mismatches 51; Indels 7; Gaps 4;

QY 133 GEAAGCGVYHSEIEEMNNVKSILHKTVDLGLP-XICNTSLYELEQLKGIAMMKELTPRTSDYLSFG 191
 DB 2 GRAVACGAPKASEIYELAVIKELHRTIDELGDSISVSGFLDEQLKGVAMMKELT 61
 QY 192 PRSDYLSFGECMSTRISAYLNKIRVAKROYDAFDIGFITTFDEFGNADILEATYPAPVA 251
 DB 62 LRTRDYLSFGECMSTRISAYLNKIRVAKROYDAFDIGFITTFDEFGNADILEATYPAPVA 121
 QY 252 KRLHGMIDDPALPVYTGFLGKMGSGAVTTLGRGSDLTATTIGKALGLREIQWKDVP 311
 DB 122 KRLHGMIDDPALPVYTGFLGKMGSCAVTTLGRGSDLTATTIGKALGLREIQWKDVP 181
 QY 312 GVLTCDPNTYPRHAKTVPEYLTPEEATLAYGAQVLAHPOSMPAREGDIPIRVKNSYNPKA 371
 DB 182 GVLTCDPNTYANVPPYLTPEEATLAYGAQVLAHPOSMPAREGDIPIRVKNSYNPKA 241
 QY 372 PGTLITRORDMDXGLVLTSLVLSKSNVTMLDIVSTRMLGQYGFARVSGICYEDLCISV 431
 DB 242 PGTVITKTRDMRKS--ILTSIVLKSNTIMLDIVSTRMLGQYGFARVSGI--FEDLGISV 297
 QY 432 DCVATSEVSVSLDPSKIMSRELIQASSELHVVHELEKIAVRLQORATISLIGNVVE 491
 DB 298 DSVATSEVSVSLDPSKIMSRELIQ--ELDHVHELEKIAVRLQORATISLIGNVQ 355
 QY 492 QSSLLIEKTRGVLRKSGVNVQMISQASKYNMSLIYHSDAKALVEALHQAFFEDDVLSQ 551
 DB 356 RSSLLIEKAFNVLRKGVNVQMISQASKYNMSLIYVNDSEAKQCVQALHSAFFENGFLSE 415
 QY 552 VEAEML 557
 DB 416 VEEADL 421

RESULT 9
 US-10-425-114-70672
 Sequence 70672, Application US/10425114
 Publication No. US20040034888A1

GENERAL INFORMATION:
 APPLICANT: Liu, Jingdong
 APPLICANT: Zhou, Yihua
 APPLICANT: Kovalic, David K.
 APPLICANT: Screen, Steven E
 APPLICANT: Tabaska, Jack E
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 FILE REFERENCE: 38-21(5313)B
 CURRENT APPLICATION NUMBER: US/10/425,114
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 73128
 SEQ ID NO 70672
 LENGTH: 426
 TYPE: PRT
 ORGANISM: *Zea mays*
 FEATURE:
 OTHER INFORMATION: Clone ID: 700343285_FLI.pep
 US-10-425-114-70672

Query Match 56.4%; Score 1587.5; DB 12; Length 426;
 Best Local Similarity 78.8%; Pred. No. 2.5e-147;
 Matches 323; Conservative 37; Mismatches 43; Indels 7; Gaps 4;

QY 145 SEIEEMNNVKSILHKTVDLGLP-XICNTSLYELEQLKGIAMMKELTPRTSDYLSFG 203
 DB 3 SEIPELAYIKDILHRTVDLGLDLSIVSGFLDEQLKGVAMMKELTPRTDYLVSFG 62
 QY 204 CMSTRISAYLNKIRVAKROYDAFDIGFITTFDEFGNADILEATYPAPVAKLHGMIDPA 263
 DB 63 CMSTRISAYLNKIRVAKROYDAFDIGFITTFDEFGNADILEATYPAPVAKLHGMIDPA 122
 QY 264 IYVYTGFLGKMGSGAVTTLGRGSDLTATTIGKALGLREIQWKDVPDVLTCDPNTYPR 323
 DB 123 IYVYTGFLGKMGSCAVTTLGRGSDLTATTIGKALGLREIQWKDVPDVLTCDPNTYAN 182
 QY 324 ACTVPEYLTPEEATLAYGAQVLAHPOSMPAREGDIPIRVKNSYNPKAPGTLITRQDMD 383
 DB 183 ALPPEYLTPEEATLAYGAQVLAHPOSMPAREGDIPIRVKNSYNPKAPGTVITKSRDMS 242
 QY 384 XGLVLTSLVLSKSNVTMLDIVSTRMLGQYGFARVSGICYEDLCISVDCVATSEVSV 443
 DB 243 KS--ILTSIVLKSNTIMLDIVSTRMLGQYGFARVSGI--FEDLGISVDCVATSEVSV 298
 QY 444 SDPSKIMSRELIQASSELHVVHELEKIAVRLQORATISLIGNVQSSLLIEKTRGV 503
 DB 299 TLDPKLSMSRELIQ--ELDHVHELEKIAVRLQORATISLIGNVQSSLLIEKAFNV 356
 QY 504 LRKSGVNVQMISQASKYNMSLIYHSDAKALVEALHQAFFEDDVLSQVE 553
 DB 357 LRKGVNVQMISQASKYNMSLIYVNDSEAKQCVQALHSAFFENGFLSE 406

RESULT 10

US-10-425-114-39779
 Sequence 39779, Application US/10425114
 Publication No. US20040034888A1

GENERAL INFORMATION:
 APPLICANT: Liu, Jingdong
 APPLICANT: Zhou, Yihua
 APPLICANT: Kovalic, David K.
 APPLICANT: Screen, Steven E
 APPLICANT: Tabaska, Jack E
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 FILE REFERENCE: 38-21(5313)B
 CURRENT APPLICATION NUMBER: US/10/425,114
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 73128
 SEQ ID NO 39779
 LENGTH: 426
 TYPE: PRT
 ORGANISM: *Glycine max*
 FEATURE:
 OTHER INFORMATION: Clone ID: 700747979_FLI.pep
 US-10-425-114-39779

Query Match 55.9%; Score 1573; DB 12; Length 426;
 Best Local Similarity 77.7%; Pred. No. 6.8e-146;
 Matches 324; Conservative 34; Mismatches 51; Indels 8; Gaps 5;

QY 140 GVLHSEIEEMNNVKSILHKTVDLGLP-XICNTSLYELEQLKGIAMMKELTPRTSDYL 198
 DB 1 GVTNVSEIEELCPFKDILHRTVDGLDVSIVAKLHELEQLKGIAMMKELTRKTDYL 60
 QY 199 VSFGECSSTRISAYLNKIRVAKROYDAFDIGFITTFDEFGNADILEATYPAPVAKLHGM 258
 DB 61 VSFGECSSTRISAYLNKIRVAKROYDAFDIGFITTFDEFGNADILEATYPAPVAKLHGM 120
 QY 259 IODPALPVYTGFLGKMGSGAVTTLGRGSDLTATTIGKALGLREIQWKDVPDVLTCDP 318
 DB 121 IODPALPVYTGFLGKMGSCAVTTLGRGSDLTATTIGKALGLREIQWKDVPDVLTCDP 180
 QY 319 NIYPHAKTVPEYLTPEEATLAYGAQVLAHPOSMPAREGDIPIRVKNSYNPKAPGTLITR 378

Db 181 NIYPKAEVPYLTFFDAALAFGAQVLAHPQSMRPAESDIPRYKNSYNPKAPGTLITK 240
 QY 379 QRDMDXGLVLTISIVKSNVTMLDIVSTRMLGQYFPLARVSGICIEDLCISVDCVATSE 438
 Db 241 ARDMSK--AVLTISIVKSNVTMLDIVSTRMLGQYFPLAKVFSI--FEEELGISTVAVATSE 296
 QY 439 VASVSLDPSKIMSRRELIOQASELDHVEELEKIAIVLLOQRAITISIGNVEOSSILE 498
 Db 297 VASVSLDPSKIMSRRELIOQ--ELDHVEELEKIAIVMLLOQRAITISIGNVEOSSILE 354
 QY 499 KTGRLRKSGVNVQMSQASKYNLSLIYHSDAVALYEAHQAFEDDVLISQVEAE 555
 Db 355 KAFRVLRITLIGITVQMSQASKYNISLVNVDSEACVRAHLIAFESE--LSELEMD 410

RESULT 11
 US-10-425-114-39716
 ; Sequence 39716; Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 39716
 ; LENGTH: 326
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700663367_FLI.pep
 ; US-10-425-114-39716

Query Match 43.5%; Score 1222.5; DB 12; Length 326;
 Best Local Similarity 79.5%; Pred. No. 1.6e-111;
 Matches 252; Conservative 24; Mismatches 34; Indels 7; Gaps 4;

QY 239 NADILEATYPAVAKRLHGMIDOPALPVYTGFLGKMKSGATTTIGRGSDLTATTICKA 298
 Db 1 NADILEATYPAVAKRLHGMIDOPALPVYTGFLGKMKSCAVTTTIGRGSDLTATTICKA 60
 QY 299 LGLREIQWKCVDGVLTCDENIYPHAKTVPYLTFFEBATELAFGAQVLAHPQSMRPAESD 358
 Db 61 LGLREIQWKCVDGVLTCDENIYPHAKTVPYLTFFEBATELAFGAQVLAHPQSMRPAESD 120
 QY 359 IPRVKNNSYNPKAPGTLITRQDMXGLVLTISIVKSNVTMLDIVSTRMLGQYFPLARV 418
 Db 121 IPRVKNNSYNPKAPGTLITRQDMXGLVLTISIVKSNVTMLDIVSTRMLGQYFPLARV 178
 QY 419 SGCIYEDLCISVDCVATSEVSVSLDPSKIMSRRELIOQASELDHVEELEKIAIVL 478
 Db 179 FSI--FEEELGISTVAVATSEVSVSLDPSKIMSRRELIOQ--ELDHVEELEKIAIVML 234
 QY 479 QQRAITISIGNVEOSSILEKTRVLRKSGVNVQMSQASKYNLSLIYHSDAVALYEA 538
 Db 235 QNSITISIGNVEOSSILEKTRVLRKSGVNVQMSQASKYNLSLVNVDSEACVRA 294
 QY 539 LHOAFEDDVLISQVEAE 555
 Db 295 LHLAFESE--LSELEMD 310

RESULT 12
 US-10-425-114-43083
 ; Sequence 43083; Application US/10425114
 ; Publication No. US20040034888A1

GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 43083
 ; LENGTH: 249
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700664751_FLI.pep
 ; US-10-425-114-43083

Query Match 35.3%; Score 992.5; DB 12; Length 249;
 Best Local Similarity 78.1%; Pred. No. 5e-89;
 Matches 193; Conservative 21; Mismatches 32; Indels 1; Gaps 1;

QY 99 RMAEVALILTFPEERPVVYISAMGKTTNNLLAGEKAVGCVTHVSEIEEMNMYSLHT 158
 Db 2 RREEVANILTSFPERPPIIVSAMGKTTNNLLAGEKAVSGVTNADSIDELNITKDLHL 61
 QY 159 KTVDELGLP-XICNTSLYELBOLKGIAMKELTPRTSDVLSFGECMSTRISAYLNKI 217
 Db 62 RTVEQLGVDRVNIENHLEBOLKGIAMKELTPRODVIYVSGECMSTRIPAYLNIL 121
 QY 218 RYKARQYADFIDGFTTDFEGNADILEATYPAVAKRLHGMIDOPALPVYTGFLGKMK 277
 Db 122 GYKARQYADFEMGIIITTDFTNADILEATYPAVAKRLHSDVSDPAIPVITGFLGKARKS 181
 QY 278 GAVTLIGRGSDLTATTIGKALGLREIQWKCVDGVLTCDENIYPHAKTVPYLTFFEBATE 337
 Db 182 CAVTLIGRGSDLTATTIGKALGLREIQWKCVDGVLTCDENIYPHAKTVPYLTFFEBATE 241

QY 338 LAYFGAQ 344
 Db 242 LAYFGAQ 248

RESULT 13
 US-10-425-114-54453
 ; Sequence 54453; Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 54453
 ; LENGTH: 221
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: UC-ZMFLMO17134C03_FLI.pep
 ; US-10-425-114-54453

Query Match 34.6%; Score 972; DB 12; Length 221;
 Best Local Similarity 95.4%; Pred. No. 4.3e-87;

Matches 207; Conservative 0; Mismatches 4; Indels 6; Gaps 3;

QY 344 QVLPQSMRPARBEDIPIRVKNSYNPKAPGTLITRQDMXGLVLTSLVLSKNVTMDI 403

Db 11 QVLPQSMRPARBEDIPIRVKNSYNPKAPGTLITRQDMXGLVLTSLVLSKNVTMDI 68

QY 404 VSTRMLGQYGFLLARVSGICIEELCTSVDCVATSEVSISVSDPSKIMRELIIQOASELD 463

Db 69 VSTRMLGQYGFLLARVSGICIEELCTSVDCVATSEVSISVSDPSKIMRELIIQOASELD 124

QY 464 HVVELEKIAIVRLLOQRAIISLIGNVEQSSIILEKTGRVLRKSGVNVQMISOGASKVM 523

Db 125 HVVELEKIAIVRLLOQRAIISLIGNVEQSSIILEKTGRVLRKSGVNVQMISOGASKVM 184

QY 524 SLIVHSDAKALVLAHQAFEDDVLSQVEAENLVG 560

Db 185 SLIVHSDAKALVLAHQAFEDDVLSQVEAENLVG 221

RESULT 14

US-09-890-813-12

; Sequence 12, Application US/09890813

; Publication No. US20020183486A1

; GENERAL INFORMATION:

; APPLICANT: E.I. du Pont de Nemours and Company

; TITLE OF INVENTION: Aspartate Kinase

; FILE REFERENCE: BRL430 PCT

; CURRENT APPLICATION NUMBER: US/09/890,813

; CURRENT FILING DATE: 2001-08-02

; PRIOR APPLICATION NUMBER: 60/172944

; PRIOR FILING DATE: 1999-12-21

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 12

; LENGTH: 261

; TYPE: PRT

; ORGANISM: Oryza sativa

US-09-890-813-12

Query Match 31.7%; Score 890.5; DB 9; Length 281;

Best Local Similarity 67.1%; Pred. No. 6.9e-79;

Matches 192; Conservative 19; Mismatches 68; Indels 7; Gaps 4;

QY 19 PPASSGHVRLAC-FGRTTGPGRAGLSMNVVADSTRRAKQADGSGVLCAPVLSGLME 77

Db 1 PPRVGRBQYLAQAAARPGRCRRRGLVRCQSGAAAVLNKDD--AASYAAAAAASS 57

QY 78 GLDQLSYVMKFGSSSVSAAARMVAAGLITFPEERPVVLSAMGKTNNLLAGEKAV 137

Db 58 ATG--PTVAMKFGSSSVASARMREVAADLISPEETPVVLSAMGKTNNLLAGEKAV 115

QY 138 GCGVIVHSEIEEMNVKSLHIKIVDELGLP-XICNTSLVLEOLKGIAMKELIPRISD 196

Db 116 SCGAPKASEIPELAVIKELHVRTIDELGLDRSIVSGLEBELQDLKGVAMMKELIPRTRD 175

QY 197 YIVSGEGCMSTRISAVYINKIRVAKROYDAFDIGFITTFDEGNADILEATYPVAKRLHG 256

Db 176 YIVSGEGCMSTRIRPAAYINKIGKARQYDAFDIGFITTFDFTNADILEATYPVAKRLHG 235

QY 257 DWIDPAIPVVTGFLGKMGSGAVTTLGRGSDLTATTIGKALGLR 302

Db 226 DWIDPAIPVVTGFLGKMGSCAVTTLGRGSDLTATTIGKALRTR 281

RESULT 15

US-10-425-114-46004

; Sequence 46004, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jindong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(5313)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 72128

SEQ ID NO 46004

LENGTH: 226

TYPE: PRT

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: 700873216_FLI.pep

US-10-425-114-46004

Query Match 27.1%; Score 762.5; DB 12; Length 226;

Best Local Similarity 73.9%; Pred. No. 2e-66; Indels 9; Gaps 4;

Matches 164; Conservative 22; Mismatches 27;

QY 335 ATELAYFGAQVLPQSMRPARBEDIPIRVKNSYNPKAPGTLITRQDMXGLVLTSLV 394

Db 1 ATELAYFGAQVLPQSMRPARBEDIPIRVKNSYNPKAPGTLITRQDMXGLVLTSLV 58

QY 395 KSNVTMDIVSTMCGYGFLLARVSGICIEELCTSVDCVATSEVSISVSDPSKIMREL 454

Db 59 KSNVTMDIVSTMCGYGFLLARVSGICIEELCTSVDCVATSEVSISVSDPSKIMREL 116

QY 455 LIOQASELDHVVELEKIAIVRLLOQRAIISLIGNVEQSSIILEKTGRVLRKSGVNVQM 514

Db 117 LIOQASELDHVVELEKIAIVRLLOQRAIISLIGNVEQSSIILEKTGRVLRKSGVNVQM 174

QY 515 SOGASKVMISLVNDSAKALVLAHQAFEDDVLSQVEAENLVG 556

Db 175 SOGASKVMISLVNDSAKALVLAHQAFEDDVLSQVEAENLVG 213

Search completed: March 23, 2004, 13:23:17

Job time : 68 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 23, 2004, 13:02:31 ; Search time 27 Seconds
(without alignments)
1995.084 Million cell updates/sec

Title: US-09-890-813-6

Perfect score: 2813
Sequence: 1 MAIPVRSAAAPRLVPSIPP.....QAFPEDVLSQVAENLLVG 560

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR 78:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1789	63.6	569	2 T48575	aspartate kinase
2	662	23.5	473	2 D64371	aspartate kinase
3	642.5	22.8	462	2 D69337	aspartate kinase
4	607	21.6	473	2 P75405	aspartate kinase
5	590	21.0	461	2 AG0452	aspartate kinase
6	569	20.2	449	1 KIECD3	aspartate kinase
7	569	20.2	449	2 C86095	aspartate kinase
8	569	20.2	449	2 G91254	lysine sensitive a
9	566	20.1	449	2 AC1013	aspartate kinase
10	531	18.9	479	2 P83328	aspartate kinase
11	516.5	18.4	519	2 T39822	aspartate kinase
12	491.5	17.5	527	1 KIRBD	aspartate kinase
13	483.5	17.2	825	2 H82086	aspartate kinase
14	477.5	17.0	820	1 DEBCK	thra difunctional
15	477.5	17.0	820	2 B85480	aspartate kinase
16	477.5	17.0	820	2 B90629	aspartate kinase
17	475.5	16.9	815	2 A64048	aspartate kinase
18	469.5	16.7	820	2 AC0502	aspartate kinase
19	468.5	16.7	820	2 F97918	aspartate kinase
20	467.5	16.6	454	2 G95047	aspartate kinase
21	464.5	16.5	828	2 G85047	aspartate kinase
22	463.5	16.5	819	2 B47057	aspartate kinase
23	461.5	16.4	916	2 T04752	aspartate kinase
24	461.5	16.4	917	2 T02954	aspartate kinase
25	455.5	16.2	819	2 AH0056	aspartate kinase
26	452.5	16.1	920	2 T02953	aspartate kinase
27	451	16.0	911	2 S46497	aspartate kinase
28	447	15.9	911	2 E86438	hypothetical prote
29	445	15.8	437	2 H96933	aspartate kinase

30	442	15.7	473	2 E72715	probable aspartate
31	441	15.7	739	2 H72364	aspartate kinase II -
32	439	15.6	915	2 T03589	probable aspartate
33	429	15.3	446	2 D90238	aspartate kinase (AKH
34	425.5	15.1	460	2 H89907	hypothetical prote
35	425	15.1	868	2 A82722	difunctional diaml
36	423.5	15.0	916	2 T06242	aspartate kinase
37	420.5	14.9	909	2 T06246	aspartate kinase
38	417.5	14.8	421	2 S15376	aspartate kinase
39	416.5	14.8	456	2 D83837	homoserine dehydro
40	415.5	14.8	421	2 S42422	aspartate kinase
41	412.5	14.7	921	2 S35160	aspartate kinase
42	409.5	14.6	421	2 F70794	probable ask prote
43	408.5	14.5	454	2 A69763	homoserine dehydro
44	408	14.5	399	2 D97123	aspartate kinase
45	402	14.3	411	2 A48946	aspartate kinase

ALIGNMENTS

RESULT 1

T48575
aspartate kinase - Arabidopsis thaliana
N/Alternate names: protein T31B5.100
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C/Accession: T48575
R/Reven, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancro
Submitted to the Protein Sequence Database, April 2000
A/Reference number: 22490
A/Accession: T48575
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-569 <BEV>
A/Cross-references: EMBL:AL163491
A/Experimental source: cultivar Columbia; BAC clone T31B5
C/genetics:
A/Map position: 5
A/introns: 134/3; 162/2; 176/1; 227/3; 277/3; 309/3; 348/3; 388/3; 419/3; 458/3; 497/
A/Note: T31B5.100

Query Match	53.6%	Score 1789;	DB 2;	Length 569;
Best Local Similarity	71.8%	Pred. No. 4.4e-115;		
Matches 369;	Conservative 54;	Mismatches 79;	Indels 12;	Gaps 6
QY	44	LSMVVADSTSRRAKQADGGDVLGAPVIGLGMGLD---	QISVYMKFGSSVSAAAR	99
DB	44	LSLPIDGDSIRKVGSGSRNIVRA-VLEEKTEATEVDEBKGITVYMKFGSSVSAAR	102	
QY	100	MAEVAGLITPEEPFPRVYVLSAMGKTNNLLAEKXVGGVHVSEIEEMNMYKSLHIK	159	
DB	103	MEVADLITPEEPFPRVYVLSAMGKTNNLLAEKXVGGVHVSEIEELSTIKELHIK	162	
QY	160	TYDEGLG-PLICNTSLVLEQLKGIAMKELTPRTSDYVSPFGCMSTRIFAYLNKIR	218	
DB	163	TYKELNIDPSVILTYLEELQGLKGIAMKELTLRTSDYVSPFGCMSTRIFAYLNKIR	222	
QY	219	VVARQYDAFDITTTDEFGNADILEATYPAVRKIGHDMVQDPAIVVMGFGKMGKSG	278	
DB	223	VVARQYDAFEIGITTTDFTMGDILEATYPAVRKELDDMHMDPAVPIVYGFGLKMGKGT	282	
QY	279	AVTTLGRGSDLTATTIGALGRLIYQWKVDVGLTCDPNIPYHAKTVPYLPFEARETEP	338	
DB	283	AVTTLGRGSDLTATTIGALGRLKIYQWKVDVGLTCDPNIPYHAKTVPYLPFEARETEP	342	
QY	339	ATFGAQLVHPQMRPARBEGDIPVRVNSYNPKAPETLITRQMDXGIVLTSTVAKSNV	398	
DB	343	ATFGAQLVHPQMRPARBEGEIPVRVNSYNPKAPETLITTRQMDTKS-LLTSTVILKRVN	400	
QY	399	TMDIVSTRMLQCYGFLARVSGICVIEDICTVDCVATSEVSVSLDPSKIMWRELIOQ	458	
DB	401	TMDIVSTRMLQCYGFLARVSGI--FEELGISTVDVATSEVSVSLDTPSKIMWRELIOQ	458	

QY 459 ASELDHVEELEKIAIVRLLOQRAIIISLIGNVEOSSILILEKTRGVLRKSGVNVOMISQGA 518
 DB 459 --ELDHVVEELEKIAIVRLLOQRAIIISLIGNVHSSILILBRAPHVLYTKGVNVOMISQGA 516
 QY 519 SKVNSLIVHSDSKALVEALHQAFFEDVLSQV 552
 DB 517 SKVNSIFVNAEAGCCVQALHKSPFESGDLSEL 550
 RESULT 2
 C64371
 aspartate kinase (BC 2.7.2.4) I - Methanococcus jannaschii
 N:Alternate names: aspartokinase I
 C:Species: Methanococcus jannaschii
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 25-Aug-2003
 C:Accession: C64371
 R:Bull, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
 Reich, C.T.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
 Iron, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
 A:Reference number: A64300; MUID:96337999; PMID:8688087
 A:Accession: C64371
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-473 <BLU>
 A:Cross-references: GB:U67506; GB:L77117; NID:g1591274; PIDN:AA98565.1; PID:g1591278; T
 C:Genetics:
 A:Map position: FOR508401-509822
 C:Superfamily: aspartate kinase; aspartate kinase homology
 C:Keywords: phosphotransferase
 F:2-468/Domain: aspartate kinase homology <DKI>
 Query Match 23.5%; Score 662; DB 2; Length 473;
 Best Local Similarity 32.3%; Pred. No. 7.2e-39;
 Matches 160; Conservative 110; Mismatches 167; Indels 58; Gaps 11;
 QY 83 LSVVKKFGSSSVSSAARMAEYAGLITFPERP--VYVLSAMGKTNNLLAGEKAVGCG 140
 DB 1 MTTWKKFGSTGVSGSERIRHVAKIYTKKKEDDDVVVVSAMSEVTALVEISQAL--D 58
 QY 141 VHVSEIEEM--NMVSLHKITVDEL-----GLPKICNTSLYEQLKKGAMKKEIT 191
 DB 59 VRDIKVGDFIKFIEKRYKAIIEKAIKSEETKEEVKIKIDSRILEKVLIGVAVGELT 118
 QY 192 PRSDYLVSGECMSTRIFSAVLRKIRKARQYDAFDIGITTFDEGNA-----DILEAT 246
 DB 119 PKSRYYILSFGERRLSSPLSGAIRLGEKS:ALGGEAGITTDNNFGSAVVKELVYKERL 178
 QY 247 YPAVAKRLHGMIDOPALPVYTGFLGKWKSGAVTTIGRGSDLTATTIGKALGLREIQV 306
 DB 179 LPL-----LKEGIIIPVYTGFLGT--TEEGYITTLGRGSDVSAALIGYGDADITEL 228
 QY 307 KCDVGVLTCDPENIYPHAKTVPYLTFEBATELAYFGAQLPQSMRAREGDIPIRVKNS 366
 DB 229 WTDVSGVYTTDPRVLPTRAPRIPKSYIAMELAIFGAKVILHPRITBEMKGIPIVKN 288
 QY 367 YNPKAPGTLITRQDMXGLVVLISVLSKSVTMLDIVSTRMLQCYGFLARVSGICTIED 426
 DB 289 FEFESEGTLLITNDMSDSIVAKISTI--KNVALINIFGAGMVGSGTAARIFKALGEEB 346
 QY 427 LGISDVCAVTSVSVSLDPSKIRWSRELIQASLELHVVELEK-----471
 DB 347 VAVILISQSEETNIS-----LVVSEEDVAKALKALRREGDGRKGSFLNNNL 394
 QY 472 IAIIVRLLOQRAIIISLIG-NVEQSSILILEKTRGVLRKSGVNVOMISQGASKVNSLIVHDS 510
 DB 395 IADVSDKDVCAIVSYGAGMGAKIGACKIFTAVSESGANTKMTAQSSEVINISFVIDEK 454
 QY 531 DAKALVEALHQAFFE 545

DB 455 DLNVCVRKLEKFE 469
 RESULT 3
 D69337
 aspartate kinase (lyec) homolog - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 25-Aug-2003
 C:Accession: D69337
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodg
 Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Uteback, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes,
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
 A:Reference number: A69250; MUID:96049343; PMID:9389475
 A:Accession: D69337
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-462 <KLE>
 A:Cross-references: GB:AE001056; GB:AE000782; NID:92689379; PIDN:AA90539.1; PID:9264;
 C:Superfamily: aspartate kinase; aspartate kinase homology
 F:2-459/Domain: aspartate kinase homology <DKI>
 Query Match 22.8%; Score 642.5; DB 2; Length 462;
 Best Local Similarity 34.4%; Pred. No. 1.6e-37;
 Matches 165; Conservative 88; Mismatches 187; Indels 39; Gaps 10;
 QY 83 LSVVKKFGSSSVSSAARMAEYAGLITFPERP--VYVYVLSAMGKTNNLLAGEKAV--G 138
 DB 1 MTTWKKFGSVSDGDNIVYHCANLVKKFSGNEVVYVVSAMQVTDALLKAKRCCSEBS 60
 QY 139 CGVI-----HVSEIE--EMNVSLHKITVDELGLPKICNTSLYEQLKKGIA 185
 DB 61 AGFIKFIADMKRKHAYEIAVAVRNDKIKAVIAAVERL-----LDELEKYLIGIS 111
 QY 186 MKKELTRISDYLVSGECMSTRIFSAVLRKIRKARQYDAFDIGITTFDEGNDILEA 245
 DB 112 YLGELTRSEDIYVSGERLSAPISFALISLVDSVALTGDAGITTRNNGRAKELPG 171
 QY 246 TVPAVAKRLHGMIDOPALPVYTGFLGKWKSGAVTTIGRGSDLTATTIGKALGLREIQ 305
 DB 172 VYTTIRNRLEPLITIKTIPVYTGFLG-ATDQSGITTLGRGSDVATALLAALDADEVW 230
 QY 306 VMKQVQDVLTCDPENIYPHAKTVPYLTFEBATELAYFGAQLPQSMRAREGDIPIRVKNS 365
 DB 221 LMKVDDIMTCDDPKYVNAVIEISYQEMELSHGAKILHPRALEPVRKNIPIRVKNS 290
 QY 366 SYNPKAPGTLITRQDMXGLVVLISVLSKSVTMLDIVSTRMLQCYGFLARVSGICTYE 425
 DB 221 TFNDAGTGYIGPTTKSSSEIVKALSLIPAGIVN--VSGAGFDPAELMSEVFGRLAE 347
 QY 426 DLTISDVCAVTSVSVSLDPSKIRWSRELIQASLELHVVELEKIAIVRLLOQRAIS 485
 DB 348 RVNIVYMAOSSSEINISLIVDIRDL--EKAYNALKELENGVYKVEKIPIT-----AVVS 399
 QY 486 LIGN-VREGSSILILEKTRGVLRKSGVNVOMISQGASKVNSLIVHDSKALVEALHQAFF 543
 DB 400 AVSGMAGTREVAKITFALGKGINVIMISQCSSEVINISFVIDSGSDGNNAKVYHDEF 458
 RESULT 4
 F75405
 aspartate kinase - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 25-Aug-2003
 C:Accession: F75405
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.C.
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uteback, T.; Zalewski, C.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: F75405
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-473 <WHI>
 A:Cross-references: GB:AE001982; GB:AE000513; NID:96459109; PIDN:AA10936.1; PID:9645911
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR1365
 A:Map position: 1
 C:Superfamily: aspartate kinase; aspartate kinase homology

Query Match 21.6%; Score 607; DB 2; Length 473;
 Best Local Similarity 33.8%; Pred. No. 5,4e-35;
 Matches 163; Conservative 96; Mismatches 185; Indels 38; Gaps 14;

```

QY 85 VVKKFGSSVSSAARMAEVAAGLITFPEE--RPVVLSAMKTTNNLLAGEKVGCGVI 142
DB 11 LVKKFGGTMMQGAHAIHRSASLAGRSVEGVVVVVSAMAGVINGLHLADAAQSGDIA 70
QY 143 HVSEIEEMMMVKSIIHKTVDLGLPICNT-----SLVE-LEQLLKGIAMKELTPRTSD 136
DB 71 RAN--DEIATLRTRFTTAQELGAAPDSETRBELREHETLRQAVGYVILRELTPRSR 128
QY 197 YVVFQECWSTRIFSAYLKIKRVKARYADAFDIFITTFDEFGNADILEATPAVAKRLHG 236
DB 129 LVVAFGERLSAPLMSIALEQSLRARIHTGSGAGILDTNFGNARPLPGTYERVGDRLSG 188
QY 257 DNIQDBAIPVVTGFLGKWKSGAVTTLGRGSDLTATTIGALGREIQWKVDVGLTC 316
DB 189 -FLSAGVTAVIAGFME-TEGATITLGGGDSFATVIGALGDEVMANKDVGWMSA 246
QY 317 DENIYFAKTVPYLTFFEEATELAYFGAQLVHPQSRPARBGDIPRVKNSYNKRAPSTLI 376
DB 247 DRRVVDGANIGVLSTGEVMEIAYFGAKVHLAVTPLQDSGIPLRVKSADPPFAGTLV 306
QY 377 TQQRMDXGLVLTSLVKSNTMDIVSTRMLQGYGLAVSGICYIEDLCISVDCVAT 436
DB 307 RQQAEEIPRPVPAVATK--NVSLITVAGAVLPEVIAVFTPAARENTTLIMVSQSS 365
QY 437 SEVSVSLSDPKIKSRELIOQA-----SELDHVEEIEKTAIVRLAQPARISLIGN- 489
DB 366 SMSNVSILAV-PSADAARAL--QALRAGLSSEIN--VEETNGV-----AVLAIVSGG 411
QY 490 VQSSSILKKTGRVLRKSGVNTQMSQASKNMGLIYHSDDAALVALHQAFFEDVL 549
DB 412 MKGQGVASARLPTALASEDINITIMTSQSSSEINISVALSGEVDPAVAHAFAF---TL 467
QY 550 SQ 551
DB 468 SQ 469
  
```

RESULT 5

AG0452
 aspartate kinase (EC 2.7.2.4) [imported] - *Yersinia pestis* (strain CO92)
 C:Species: *Yersinia pestis*
 C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 25-Aug-2003
 C:Accession: AG0452
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarrega, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.;
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11566360
 A:Accession: AG0452
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-461 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC93187.1; PID:915981636; GSPDB:GN00175
 C:Genetics:
 A:Gene: lysC
 C:Superfamily: aspartate kinase; aspartate kinase homology

C:Keywords: phosphotransferase

Query Match 21.0%; Score 590; DB 2; Length 461;
 Best Local Similarity 34.5%; Pred. No. 8,1e-34;
 Matches 165; Conservative 92; Mismatches 173; Indels 48; Gaps 15;

```

QY 84 SVVKKFGSSVSSAARMAEVAAGLITFPEEPPVVLSAMKTTNNLLA-----GEKAV 137
DB 16 TVVAKFGGTVAASFDASRSADVLSNPDLVILSASAGIT--NLVLVADDSBEKKA 73
QY 138 GCGVTHVEIE--EMMVKSIIHKTVDLGLPICNTSLVELBQLKGIAMKE-----L 190
DB 74 D---HEDIRHIQYDIARTDSTVIR-----EIDRMENIMLSEASLST 118
QY 191 TPRSDYVVFQECWSTRIFSAYLKIKRVKARYADAFDIFITTFDEFGNADILEATPAV 250
DB 119 SPALDELVSNGELMSTLTFELLRQVAVEMVDVAKV--MRTNDRGRAEPDTSALAE 177
QY 251 AKRLHDMIQDBAIPVVTGFLGKWKSGAVTTLGRGSDLTATTIGALGREIQWKV 310
DB 178 AQTLLAPRIED-ATVVTQGFISGK--GRITLGRGSDYVALLGLALVSRIDWTDV 235
QY 311 DGVLTCDENIYFAKTVPYLTFFEEATELAYFGAQLVHPQSRPARBGDIPRVKNSYNK 370
DB 236 PGITTDPRVYPAKRIKIFEEAEMATFGAKIHPATLVAVSDIMFVGSSKDP 295
QY 371 AEGTLITQRMDXGLVLTSLVKSNTMDIVSTRMLQGYGLAVSGICYIEDLCIS 430
DB 296 AGGLTVLQNE---TYNPLFRALARRKQTLTLTSLMMLHARFLEAVFNI--LASHSIS 350
QY 431 VDCVATSEVSLSDPKIKSRELIOQA-----SELDHVEEIEKTAIVRLAQPARISLIG 488
DB 351 VDLITSEVSVALTIDTIGTSTGDSLTITS-----LTLESSICRVEVEDLALVAII 405
QY 489 -NVQSSSILKKTGRVLRKSGVNTQMSQASKNMGLIYHSDDAALVALHQAFF 545
DB 406 NNLQAGVAGKEVFGVL--DFNIRMI CYGASHNLCELVPGNDADKRVQTLHYNDF 461
  
```

RESULT 6

K12CD3
 aspartate kinase (EC 2.7.2.4) III, lysine-sensitive - *Escherichia coli* (strain K-12)
 N:Alternate names: aspartokinase III
 C:Species: *Escherichia coli*
 C>Date: 31-Mar-1988 #sequence_revision 17-Oct-1997 #text_change 25-Aug-2003
 C:Accession: G65209; A25659; T41098
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Berna, N.T.; Burland, V.; Riley, M.;
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of *Escherichia coli* K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: G65209
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-449 <BLAT>
 A:Cross-references: GB:AE000475; GB:U00096; NID:91790448; PIDN:AACT6994.1; PID:917904
 A:Experimental source: strain K-12, substrain MG1655
 R:Casas, M.; Parsot, C.; Cohen, G.N.; Patte, J.C.
 J. Biol. Chem. 261, 1052-1057, 1986
 A:Title: Nucleotide sequence of *lysC* gene encoding the lysine-sensitive aspartokinase
 A:Reference number: A25659; MUID:86111734; PMID:3003049
 A:Accession: A25659
 A:Molecule type: DNA
 A:Residues: 1-57; C, 59-400, A, 402-449 <CAS>
 A:Cross-references: GB:M1812; NID:9146682; PIDN:AAA24095.1; PID:9146683
 A:Experimental source: K12
 R:Casas, M.; Ronceray, J.; Patte, J.C.
 Nucleic Acids Res. 11, 6157-6166, 1983
 A:Title: Nucleotide sequence of the promoter region of the *E. coli lysC* gene.
 A:Reference number: 141098; MUID:84015362; PMID:6312411
 A:Accession: 141098
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-13,'AS',16-19,'E',<RES>
 A:Cross-references: EMBL:X00008; NID:G41937; PIDD:CA24910.1; PID:G41938
 C:Comment: The active enzyme is a dimer of identical chains, unlike the other two aspartate kinases, which is one of the end metabolites of the aspartic acid family branched C:Genetics:
 A:Gene: lysc
 A:Map position: 91 min
 C:Superfamily: aspartate kinase; aspartate kinase homology
 C:Keywords: ATP; lysine biosynthesis; phosphotransferase
 F:4-448/Domain: aspartate kinase homology <DK1>

Query Match 20.2%; Score 569; DB 1; Length 449;
 Best Local Similarity 34.0%; Pred. No. 2,4e-32;
 Matches 161; Conservative 89; Mismatches 182; Indels 42; Gaps 14;
 QY 85 VVKKFGSSVSSAARMAEVAAGLITPEEPFVVVLSAMKTTNNLLAGEKAVGGCVTHV 144
 DB 5 VVSKFGSTSVADFPAMRSADIVLSDANVR-LVVLASASAGITNLVALAE-----GLEPG 58
 QY 145 SEIEEMNVKSLHKTVDLGLPYICNTSLYELEQLKGIAMKE-----LTPRTSDLYV 199
 DB 59 ERPEKLDLRNIQALIERLRYP---NVIREIERLENTIYLAEMALATSPALTDLY 115
 QY 200 SFGECSMTRIFSAYLANKIRYKARYDAFDIGFTTDFGNAD-----LLEATYPAVAKRL 254
 DB 116 SHGELMSTLLEVEILREEDVOQWDFVRKV-MRTNDRFGAEEDIALAEALALQLEPRL 174
 QY 255 HGDWIDDPALPVVTGFLGKMGSGAVTTLGRGSDTLATTIGALGLREIQVWKDVGVL 314
 DB 175 -----NEGAVITGFGISENK-GRITTLGRGSDTYALLAEALHASVDIWTDPGITY 227
 QY 315 TCDPNIYPAKTVPLTPEEATELAYGAVLHPQSMRPARBGDIPIRVKNSYNKAPGT 374
 DB 228 TTDPRVSAKRIDEIIFAEAEKATPGACVLPATLLPAVSDIPIVFGSSKDPRAAGT 287
 QY 375 LITRDMDXGLVVLTSIVLKSNTMDIVSTFMLQYGFARVSGICYEDLCISYDCV 434
 DB 288 LVCKNTE---NPPLFRLALRRQTLTLHSLMLSRGFLAEVGIILARHN--ISVDLI 342
 QY 435 ATSEVSVSLDPEKINSR--ELIQASELDHVELEKIAVRLQCAITISLIGN-VE 491
 DB 343 TTSEVSVALTDDTGSTGSDTLTOS-----LIMELSLCREVEEGALVALIGNDLS 397
 QY 492 QSLILEKTRGVLRKSGVNVQMTSGASKVMSLIYHSDAKALYEALHQAFFE 545
 DB 398 KACGVGKEVGVLEP--FNIRMICYGASSHNLCELVPGEDAQVQKLSNLF 449

RESULT 7

aspartokinase III, lysine sensitive [imported] - Escherichia coli (strain O157:H7, sub
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 25-Aug-2003
 C:Accession: C86095
 R:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouelis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: C86095
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-449 <STO>
 A:Cross-references: GB:AE005174; NID:G12518967; PIDD:AGS9223.1; GSPDB:GN00145; UMDP:Z56
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: lysc
 C:Superfamily: aspartate kinase; aspartate kinase homology

Query Match 20.2%; Score 569; DB 2; Length 449;
 Best Local Similarity 34.2%; Pred. No. 2,4e-32;
 Matches 162; Conservative 89; Mismatches 182; Indels 42; Gaps 14;

QY 85 VVKKFGSSVSSAARMAEVAAGLITPEEPFVVVLSAMKTTNNLLAGEKAVGGCVTHV 144
 DB 5 VVSKFGSTSVADFPAMRSADIVLSDANVR-LVVLASASAGITNLVALAE-----GLEPG 58
 QY 145 SEIEEMNVKSLHKTVDLGLPYICNTSLYELEQLKGIAMKE-----LTPRTSDLYV 199
 DB 59 ERPEKLDLRNIQALIERLRYP---NVIREIERLENTIYLAEMALATSPALTDLY 115
 QY 200 SFGECSMTRIFSAYLANKIRYKARYDAFDIGFTTDFGNADLLEATYPAVAKRLHGM 259
 DB 116 SHGELMSTLLEVEILREEDVOQWDFVRKV-MRTNDRFGAE-----EDVALAEAL 168
 QY 260 Q-----DPAIPVVTGFLGKMGSGAVTTLGRGSDTLATTIGALGLREIQVWKDVGVL 314
 DB 169 QLLPRLNGVITIGFGISENK-GRITTLGRGSDTYALLAEALHASVDIWTDPGITY 227
 QY 315 TCDPNIYPAKTVPLTPEEATELAYGAVLHPQSMRPARBGDIPIRVKNSYNKAPGT 374
 DB 228 TTDPRVSAKRIDEIIFAEAEKATPGACVLPATLLPAVSDIPIVFGSSKDPRAAGT 287
 QY 375 LITRDMDXGLVVLTSIVLKSNTMDIVSTFMLQYGFARVSGICYEDLCISYDCV 434
 DB 288 LVCKNTE---NPPLFRLALRRQTLTLHSLMLSRGFLAEVGIILARHN--ISVDLI 342
 QY 435 ATSEVSVSLDPEKINSR--ELIQASELDHVELEKIAVRLQCAITISLIGN-VE 491
 DB 343 TTSEVSVALTDDTGSTGSDTLTOS-----LIMELSLCREVEEGALVALIGNDLS 397
 QY 492 QSLILEKTRGVLRKSGVNVQMTSGASKVMSLIYHSDAKALYEALHQAFFE 545
 DB 398 KACGVGKEVGVLEP--FNIRMICYGASSHNLCELVPGEDAQVQKLSNLF 449

RESULT 8

lysine sensitive aspartokinase III [imported] - Escherichia coli (strain O157:H7, sub
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 25-Aug-2003
 C:Accession: G91254
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurikawa, K.; Ishii, K.; Yokoyama, K.; Han, C
 Gatawara, N.; Yasunaga, T.; Kihara, S.; Shida, T.; Hattori, M.; Shingawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: G91254
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-449 <HAY>
 A:Cross-references: GB:BA000007; PIDD:BA038430.1; PID:G13364483; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: ECS507
 C:Superfamily: aspartate kinase; aspartate kinase homology

Query Match 20.2%; Score 569; DB 2; Length 449;
 Best Local Similarity 34.2%; Pred. No. 2,4e-32;
 Matches 162; Conservative 88; Mismatches 182; Indels 42; Gaps 14;
 QY 85 VVKKFGSSVSSAARMAEVAAGLITPEEPFVVVLSAMKTTNNLLAGEKAVGGCVTHV 144
 DB 5 VVSKFGSTSVADFPAMRSADIVLSDANVR-LVVLASASAGITNLVALAE-----GLEPG 58
 QY 145 SEIEEMNVKSLHKTVDLGLPYICNTSLYELEQLKGIAMKE-----LTPRTSDLYV 199
 DB 59 ERPEKLDLRNIQALIERLRYP---NVIREIERLENTIYLAEMALATSPALTDLY 115
 QY 200 SFGECSMTRIFSAYLANKIRYKARYDAFDIGFTTDFGNADLLEATYPAVAKRLHGM 259
 DB 116 SHGELMSTLLEVEILREEDVOQWDFVRKV-MRTNDRFGAE-----PDVALAEAL 168
 QY 260 Q-----DPAIPVVTGFLGKMGSGAVTTLGRGSDTLATTIGALGLREIQVWKDVGVL 314
 DB 169 QLLPRLNGVITIGFGISENK-GRITTLGRGSDTYALLAEALHASVDIWTDPGITY 227

```

QY 315 TCDPNITYPAKTVPYLTREBEATELAYFGAQLVHPQSMRPAEGDIPIRVKNSYNPKAPGT 374
DB 228 TTDPRVSAKRIIDIAEFAEAEMATFPAKYLHPATLLPAVRSDIPFVFGSKDPRAGGT 287
QY 375 LTRQDMXGLVLTSLVLSKSNVTMLDIVSRMIGQVGFARFSGICYEDLCISYDCV 434
DB 288 LVCKNKE--NPFRLALRNQTLTLHSLNMLHSGFLAEVFGILARN--ISVDLI 342
QY 435 ATSEVSVSLDPSKIWSR--ELIQASELDHVEELEKIALVRLQCPATISLIG-VE 491
DB 343 TTSEVSVSLDPSKIWSR--ELIQASELDHVEELEKIALVRLQCPATISLIG-VE 397
QY 492 QSLILEKTRVLRKSGVNVOMISOGASKNNMSLIVHSDPAKALVEALHQAFFE 545
DB 398 KACGVKGVLEP--FNIRMICYGASHNLCFLVPGSDAEQVVKLHSLNFE 449

RESULT 9
AC1013
aspartate kinase (EC 2.7.2.4) - Salmonella enterica subsp. enterica serovar Typh1 (strai
C/Spectes: Salmonella enterica subsp. enterica serovar Typh1
A/Note: this species has also been called Salmonella typh
C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 25-Aug-2003
C/Accession: AC1013
R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A/Reference number: AB0502; MUID:21534947; PMID:11677608
A/Accession: AC1013
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-449 <PAR>
A/Cross-references: GB:AL51382; PIDN:CAD09204.1; PID:916505208; GSPDB:GN00176
C/Genetics:
A/Gene: STY4416
C/Superfamily: aspartate kinase; aspartate kinase homology
C/Keywords: phosphotransferase

Query Match 20.1%; Score 566; DB 2; Length 449;
Best Local Similarity 34.2%; Pred. No. 3.8e-32;
Matches 162; Conservative 85; Mismatches 185; Indels 42; Gaps 14;

QY 85 VNMFGSSVSSAARAEVAGLITFPERPVPVVSAMGKTTNNLLAGEKAVGCVTHV 144
DB 5 VAKFGGTIVADPDAMNRSDIYLSANR-LVLSASAGITNLVALAE-----GMEPG 58
QY 145 SEIEEMNVKSLHKTVDLGLPXICNTSLYELBOLKGIAMKE-----LTPKSDLV 199
DB 59 ERRATDAIRKIOFDLIDRLHP--NVIREIERLENTITLAEASLATSALTDELV 115
QY 200 SPFGCWSTRIFSAVINKIRKARQYDAFDIGITTFDEFGNAD-----ILEATYPAVARL 254
DB 116 SHGELNSTLFLVEILRQDYQAMFVDRK-METSRFGAREPDAALAEALAQQLRL 174
QY 255 HGDWIDPAIPVYTGFLGKMGKGAVTTLGRGSDLTATTIGKALGREIQWQVQVGL 314
DB 175 -----SETLVITQGFISSEK-GRITTLGRGSDTYALALAEALMAVNDITWTVPGIY 227
QY 315 TCDPNITYPAKTVPYLTREBEATELAYFGAQLVHPQSMRPAEGDIPIRVKNSYNPKAPGT 374
DB 228 TTDPRVSAKRIIDIAEFAEAEMATFPAKYLHPATLLPAVRSDIPFVFGSKDQAGGT 287
QY 375 LTRQDMXGLVLTSLVLSKSNVTMLDIVSRMIGQVGFARFSGICYEDLCISYDCV 434
DB 288 LVCKNKE--NPFRLALRNQTLTLHSLNMLHSGFLAEVFGILARN--ISVDLI 342
QY 435 ATSEVSVSLDPSKIWSR--ELIQASELDHVEELEKIALVRLQCPATISLIG-VE 491
DB 398 KACGVKGVLEP--FNIRMICYGASHNLCFLVPGSDAEQVVKLHSLNFE 449

```

```

QY 492 QSLILEKTRVLRKSGVNVOMISOGASKNNMSLIVHSDPAKALVEALHQAFFE 545
DB 398 KACGVKGVLEP--FNIRMICYGASHNLCFLVPGSDAEQVVKLHSLNFE 449

RESULT 10
F82328
aspartate kinase III, lysine-sensitive VC0391 [imported] - Vibrio cholerae (strain N1696
C/Spectes: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 25-Aug-2003
C/Accession: F82328
R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gmin, M.L.; Dodson, R.
chardeon, D.; Esmolaeva, M.D.; Vamathevan, J.; Bae, S.; Qin, H.; Dragol, I.; Sellers
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A/Reference number: AB2035; MUID:20406833; PMID:10952301
A/Accession: F82328
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-479 <HEI>
A/Cross-references: GB:AB004127; GB:AB003852; MID:9654808; PIDN:AAF93564.1; GSPDB:GN
A/Experimental source: serogroup O1; strain N16961; biotype El Tor
C/Genetics:
A/Gene: VC0391
A/Map position: 1
C/Superfamily: aspartate kinase; aspartate kinase homology

Query Match 18.9%; Score 531; DB 2; Length 479;
Best Local Similarity 32.6%; Pred. No. 1.2e-29;
Matches 155; Conservative 95; Mismatches 183; Indels 42; Gaps 16;

QY 86 VNMFGSSVSSAARAEVAGLITFPERPVPVVSAMGKTTNNLLAGEKAVGCVTHV 145
DB 34 VAKFGGTIVADPDAMNRSDIYLSANR-LVLSASAGITNLVALAE-----GMEPG 58
QY 146 EIEE-----NMVKSILHKTVDLGLPXICNTSLYELBOLKGIAMKE-----LTPKSD 196
DB 85 STTRQALIQADLHVIYIIDLADPTVER--EVRKIVDSVSAEASFOSSKULTD 141
QY 197 YVSGECWSTRIFSAVINKIRKARQYDAFDIGITTFDEFGNAD-----ILEATYPAVARL 254
DB 142 HVAAGELISTYITLQWREKRGVDAFREDIEV-LRTDHDGRAEPQLKETAOLAKELV 200
QY 255 HGDWIDPAIPVYTGFLGKMGKGAVTTLGRGSDLTATTIGKALGREIQWQVQVGL 314
DB 201 P--LCQYVVVVGITGSD-AQNTTTLGRGSDYSALAEAVQASGLEITWTVPGIY 256
QY 315 TCDPNITYPAKTVPYLTREBEATELAYFGAQLVHPQSMRPAEGDIPIRVKNSYNPKAPGT 374
DB 257 STDPRIAPKAAPIPEISFEBAEYANFGAKLHPETLLPAVRHGIPIVSGSKPEKGT 316
QY 375 LTRQDMXGLVLTSLVLSKSNVTMLDIVSRMIGQVGFARFSGICYEDLCISYDCV 434
DB 317 WIRKQVE--SAVLRLALRNQTVTLRSQWHAAGFLAKFEL--LAKHKTIVDLI 371
QY 435 ATSEVSVSLDPSKIWSR--ELIQASELDHVEELEKIALVRLQCPATISLIG-VEQ 492
DB 372 TTSEIVSLTDLKT-----DTGGAPLPEAFAVRLAEELCEVEYVNCLIALIGNKKE 426
QY 493 QSLILEKTRVLRKSGVNVOMISOGASKNNMSLIVHSDPAKALVEALHQAFFE 547
DB 427 SRGVAKQVSVL--EDYNLRMICYGASHNLCFLHESVSKQAVKHLKELEFEB 479

RESULT 11
T39822
aspartate kinase - fission yeast (Schizosaccharomyces pombe)
C/Spectes: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000
C/Accession: T39822
R/Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Bothe, G.; Fohl, T.

```


Nature 406, 477-483, 2000
A>Title: DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A/Reference number: AB2035; MUID:20406833; PMID:10953301
A/Accession: H82086
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1825 <H1>
A/Cross-references: GB:AE004306; GB:AE003852; NID:g9656924; PIDN:AA95507.1; GSPDB:GN001
A/Experimental source: serogroup O1, strain N16961; Diotype E1 Tor
C/Genetics:
A/Map position: 1
A/Superfamily: bifunctional aspartokinase/homoserine dehydrogenase I; aspartate kinase H

Query Match 17.2%; Score 483.5; DB 2; Length 825;
Best Local Similarity 29.0%; Pred. No. 5.9e-26;
Matches 137; Conservative 96; Mismatches 211; Indels 29; Gaps 11;

86 VMKFGSSVSSAARMAEVAAGLITFEERPV-VLISAMGKTNNLLAGEKAVGCVH 144
9 VIKFGSSLADPERPLRADIIANNAQSEVAVLSAPKTKNLVAATESALRQGEVA 68
145 SEIEBNNVKSLS-----HKTVDGLPXCINTSLYELQLLKGLAMKELTPRTDYL 198
69 QIVLESSFYALLDGIKALPNDSDAYQQOVHSSMTQRFVHGITLLGMCPDNVARI 128
199 VSFGECSSTRIFSAVYLNKIRVKARQYDAFDIGITTFDEFGNDILEATVPA--VAKRLHG 256
129 ISKGERVSIOLMKAAMWAKLPANLYD--PVKTL-----LAKGDHLEAMVDVLSIQRRQ 182
257 DWIQDPAIPVATGFLGKMKSGAVTTLGRGSDLTATTGKALGREIQVKDVGVLTC 316
183 APIPQGHVNIIMPETA-GNAQGLVLCGRGSDYSAAVLAACIRADCCETWDDVGVNVC 241
317 DPNYIPAKTVPTLFEATEALVAPGQVLPHPQSMAPREGDIPYRVKSVYKAPGTII 376
242 DPRVDDARLKLKISYSEANELSYFGASVHPETIPIAFOFPLIKNSFPQAGTII 301
377 TQRDMXGLVLTISVLSKSNVMTLDIVSTRMLGOYGLFARVSGICYEIDLCISVDCAT 436
302 GQDTGEK--LAIKGIITLISNLIMVNVSGPMKMGVMAKRVGAMSAVDVIVITQSS 359
437 SEVSVSVDLP--KIMRELIQASL---DHVELEKALVRLQLQAAIISLGN--VE 491
360 SEYSISFCIEAQAALADAFELKXGLLEPEFVDN-----AIIITLVGDGR 413
492 QSSILTEKTRVLRKSGVNVOMISQASKYNMGLIVHSDAKALVEALHQAFF 544
414 TSKVSAQFFSSLAHVNVVIALAQSSSEALISAVIPDKISEALVACHENLF 466

RESULT 14
DEBCK
Chia bifunctional enzyme [validated] - *Escherichia coli* (strain K-12)
N/Alternate names: aspartokinase I / homoserine dehydrogenase I; protein f132
N/Contents: aspartate kinase (EC 2.7.2.4) I [validated]; homoserine dehydrogenase (EC 1.3.5.3)
C/Species: *Escherichia coli*
C/Date: 31-Dec-1980 #sequence revision 10-Oct-1997 #ext change 18-Aug-2003
C/Accession: B64720; S56629; A00671; A15659; A14560; S40531; I57719; I69770; I73480
R/Battner, F.R.; Plunkett III, G.; Bloch, C.A.; Perina, N.T.; Burland, V.; Riley, M.; Co
A./Rose, D.J.; Mau, B.; Shaoh, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of *Escherichia coli* K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: B64720
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-820 <BLAT>
A/Cross-references: GB:AE000111; GB:U00096; NID:g1786181; PIDN:AACT3113.1; PID:g1786183;
A/Experimental source: strain K-12, substrain MG1655
R/Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.
Nucleic Acids Res. 25, 2105-2119, 1995
A>Title: Analysis of the *Escherichia coli* genome VI: DNA sequence of the region from 92

A/Reference number: S56314; MUID:95334362; PMID:7610040
A/Accession: S56629
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-820 <BLAT>
A/Cross-references: EMBL:U14003; NID:g1263172; PIDN:AAA97301.1; PID:g537245
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994
R/Cosart, P.
unpublished results, cited by Zakin, M.M., Duchange, N., Ferrara, P., and Cohen, G.N.
A/Reference number: A00671
A/Accession: A00671
A/Molecule type: DNA
A/Residues: 1-229, 'N', 231-374, 'L', 376-392, 'A', 394-405, 'L', 407-552, 'N', 554-606, 'T', 606
R/Katinka, M.; Cosart, P.; Sibilli, L.; Saint-Giron, I.; Chalvignac, M.A.; Le Bras,
Proc. Natl. Acad. Sci. U.S.A. 77, 5730-5733, 1980
A>Title: Nucleotide sequence of the *thrA* gene of *Escherichia coli*.
A/Reference number: A15659; MUID:81077247; PMID:7003595
A/Accession: A15659
A/Molecule type: DNA
A/Residues: 1-10, 'L', 12-229, 'N', 231-374, 'L', 376-392, 'A', 394-405, 'L', 407-552, 'N', 554-606
A/Cross-references: GB:J01706; EMBL:V00360; NID:g147977; PIDN:AAA83914.1; PID:g147975
A/Note: this sequence has been revised in reference A00671
R/Sibilli, L.; Le Bras, G.; Cosart, P.; Chalvignac, M.A.; Le Bras, G.; Briley, P.A.;
Biochimie 61, 723-729, 1979
A>Title: The primary structure of *Escherichia coli* K 12 aspartokinase I-homoserine de
A/Reference number: A14560; MUID:80043179; PMID:387092
A/Accession: A14560
A/Molecule type: protein
A/Residues: 51-112, 'E', 114-129, <SIB>
R/Yara, T.; Mori, H.; Nagai, H.; Nagata, T.; Ishihama, A.; Fujita, N.; Isoo, K.; Mit
submitted to the EMBL Data Library, December 1992
A/Description: Systematic sequencing of the *Escherichia coli* genome: analysis of the
A/Reference number: S40531
A/Accession: S40531
A/Molecule type: DNA
A/Residues: 1-10, 'L', 12-229, 'N', 231-374, 'L', 376-392, 'A', 394-405, 'L', 407-552, 'N', 554-606
A/Cross-references: EMBL:D10483; NID:g216434; PIDN:AAA01286.1; PID:g216435
A/Experimental source: strain K-12
R/Cosart, P.; Katinka, M.; Yaniv, M.
Mol. Gen. Genet. 175, 39-44, 1979
A>Title: Construction and expression of a hybrid plasmid containing the *Escherichia*
A/Reference number: I57719; MUID:80077291; PMID:390305
A/Accession: I57719
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 553-586, 'TT' <COS2>
A/Cross-references: GB:M10644; NID:g147982; PIDN:AAA24671.1; PID:g147983
R/Gardner, J.F.
J. Biol. Chem. 257, 3886-3904, 1982
A>Title: Initiation, pausing, and termination of transcription in the threonine oper
A/Reference number: I55222; MUID:82142573; PMID:6277952
A/Accession: I69970
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-10 <GR>
A/Cross-references: EMBL:X68872; NID:g43073; PIDN:CAA48734.1; PID:g43075
R/Lynn, S.P.; Bauer, C.E.; Chapman, K.A.; Gardner, J.F.
J. Mol. Biol. 183, 529-541, 1985
A>Title: Identification and characterization of mutants affecting transcription term
A/Reference number: I56399; MUID:85264808; PMID:2410621
A/Accession: I73480
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-10 <LVN>
A/Cross-references: GB:M28570; NID:g290476; PIDN:AAA24673.1; PID:g290478
C/Genetics:
A/Map position: 0 min
A/Function: <ASP>
A/Description: as aspartate kinase, catalyzes the phosphorylation by ATP of aspartat
A/Pathway: aspartate metabolism
A/Note: regulated allosterically by L-threonine
C/Function: <HOM>

A:Description: as homoserine dehydrogenase catalyzes the reduction by NADH of aspartate
 A:Pathway: glycine metabolism; serine metabolism; threonine biosynthesis
 C:Superfamily: bifunctional aspartokinase/homoserine dehydrogenase I; aspartate kinase
 C:Keywords: allosteric regulation; homodimer; multifunctional enzyme; oxidoreductase;
 F1-460/Domain: aspartate kinase homology <DK1>
 F1-463-719/Domain: homoserine dehydrogenase homology <HSD>
 F1-467-495/Region: beta-alpha-beta NAD(P) nucleotide-binding fold

Query Match 17.0%; Score 477.5; DB 1; Length 820;
 Best Local Similarity 29.4%; Pred. No. 1.5e-25;
 Matches 142; Conservative 98; Mismatches 210; Indels 33; Gaps 13;

```

QY 86 VMKFGSSVSSAARMAEVAAGLITPPEERPV-VILSAMGKTNLLAGEKAVG--CGVI 142
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   3 VLKFGGTSVANAERFLRVADILSNAROGVATVLSAPAKITNHLVAMIEKTIISGQDALP 62
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   143 HVSEIEEMNVKSLHITVDELGLP-----XICNTSLYELBOLKGIAMMKELTPRTSDY 197
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 63 NISDAER-IFAEILITGLAAAPGFPPLAQKTVDQEPQIKVHLHGISLGGCPDSINAA 121
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   QY 198 LVSFGCMSTRIFSAVINKIRVKAROYDAFDIG---FITDEF--GNADILEATYPAVA 251
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 122 LICRGKXSIAMAGVL-----EARGHNVTVIDPEVKLLAVGHYLESTVDIAESTRIAA 176
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   QY 252 KRLHGDWIDPAIPVVTGFLGKMGKSGAVTTIGRGSDLTATTIGALGRLQVWKD 311
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 177 SRIPAD-----HNVIMAGFTA-GNEKGELVVLGRNSDYSAAVLAACLRADCCETWTDV 230
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   QY 312 GVLTCDPNITYPHAKTVPLTPEEATELAFGAQVLPQSMRPARAGDIPRVNSYNPKA 371
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 231 GYTTCDPQVDPDAILKMSYQGMELSTFGAKVLRHTITPIAQFOIPLIKNTGNPQA 290
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 122 LICRGKXSIAMAGVL-----EARGHNVTVIDPEVKLLAVGHYLESTVDIAESTRIAA 176
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   QY 252 KRLHGDWIDPAIPVVTGFLGKMGKSGAVTTIGRGSDLTATTIGALGRLQVWKD 311
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 177 SRIPAD-----HNVIMAGFTA-GNEKGELVVLGRNSDYSAAVLAACLRADCCETWTDV 230
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   QY 312 GVLTCDPNITYPHAKTVPLTPEEATELAFGAQVLPQSMRPARAGDIPRVNSYNPKA 371
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 231 GYTTCDPQVDPDAILKMSYQGMELSTFGAKVLRHTITPIAQFOIPLIKNTGNPQA 290
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   QY 372 PGTLLTRQDMXGLVLTSLVLSKNTMLDIVSTRMLGQYGLFARYSGICYIEDLCISV 431
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 291 PGTLLIGASRDEDE--LPVKGISLNMMAMFVSFGPMKGMVGMARVFAAMSRARISVYL 348
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   QY 432 DCVATSEVSVSLSDPK-IMSRELIQASSELDHVVELEKIAIVRLQORATISLIGN- 489
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 349 ITQSSSEYISFCVPSQDCVAREAMQEFYLEKEGLLEPLAVERL---AIIISVGDG 405
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   QY 490 VEQSSILIEKTRGVLRKSGVNVQMTSGASKVNMGLIYHSDAKALVEALHQAFFEDVL 549
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 406 MRTLRGISAKFFALAPANINIVAIQSSERSISVVVNNDDATTGVRVTHQMLFNTDOV 465
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   QY 550 SQV 552
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 466 IEV 468
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

RESULT 15

B85480
 aspartokinase I, homoserine dehydrogenase I [imported] - Escherichia coli (strain O157:H7)
 C/Species: Escherichia coli
 C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 18-Aug-2003
 C/Accession: B85480
 R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A/Reference number: A85480; MUID:21074935; PMID:11206551
 A/Accession: B85480
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-820 <STO>
 A/Cross-references: GB:AE005174; NID:G12512677; PIDN:AGS4302.1; GSPDB:GN00145; UWGP:Z00
 A/Experimental source: strain O157:H7, substrain EDL933
 C/Genetics:
 A/Gene: thra
 C:Superfamily: bifunctional aspartokinase/homoserine dehydrogenase I; aspartate kinase I

Query Match 17.0%; Score 477.5; DB 2; Length 820;
 Best Local Similarity 29.4%; Pred. No. 1.5e-25;
 Matches 142; Conservative 98; Mismatches 210; Indels 33; Gaps 13;

```

QY 86 VMKFGSSVSSAARMAEVAAGLITPPEERPV-VILSAMGKTNLLAGEKAVG--CGVI 142
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   3 VLKFGGTSVANAERFLRVADILSNAROGVATVLSAPAKITNHLVAMIEKTIISGQDALP 62
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   QY 143 HVSEIEEMNVKSLHITVDELGLP-----XICNTSLYELBOLKGIAMMKELTPRTSDY 197
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 63 NISDAER-IFAEILITGLAAAPGFPPLAQKTVDQEPQIKVHLHGISLGGCPDSINAA 121
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   QY 198 LVSFGCMSTRIFSAVINKIRVKAROYDAFDIG---FITDEF--GNADILEATYPAVA 251
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 122 LICRGKXSIAMAGVL-----EARGHNVTVIDPEVKLLAVGHYLESTVDIAESTRIAA 176
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   QY 252 KRLHGDWIDPAIPVVTGFLGKMGKSGAVTTIGRGSDLTATTIGALGRLQVWKD 311
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 177 SRIPAD-----HNVIMAGFTA-GNEKGELVVLGRNSDYSAAVLAACLRADCCETWTDV 230
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   QY 312 GVLTCDPNITYPHAKTVPLTPEEATELAFGAQVLPQSMRPARAGDIPRVNSYNPKA 371
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 231 GYTTCDPQVDPDAILKMSYQGMELSTFGAKVLRHTITPIAQFOIPLIKNTGNPQA 290
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   QY 372 PGTLLTRQDMXGLVLTSLVLSKNTMLDIVSTRMLGQYGLFARYSGICYIEDLCISV 431
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 291 PGTLLIGASRDEDE--LPVKGISLNMMAMFVSFGPMKGMVGMARVFAAMSRARISVYL 348
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   QY 432 DCVATSEVSVSLSDPK-IMSRELIQASSELDHVVELEKIAIVRLQORATISLIGN- 489
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 349 ITQSSSEYISFCVPSQDCVAREAMQEFYLEKEGLLEPLAVERL---AIIISVGDG 405
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   QY 490 VEQSSILIEKTRGVLRKSGVNVQMTSGASKVNMGLIYHSDAKALVEALHQAFFEDVL 549
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 406 MRTLRGISAKFFALAPANINIVAIQSSERSISVVVNNDDATTGVRVTHQMLFNTDOV 465
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   QY 550 SQV 552
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
   DB 466 IEV 468
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

Search completed: March 23, 2004, 13:18:04
 Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 11:07:15 ; Search time 19 Seconds
(without alignments)
1534.699 Million cell updates/sec

Title: US-09-890-813-6

Perfect score: 2813
Sequence: 1 MALPVSAAAPRLVPSIRP.....QAFEDVLSQVEANLLVG 560

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt 42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	662	23.5	AK_METUA	Q57991 mechanococ
2	569	20.2	AK3_ECOI	P08660 escherichia
3	516.5	18.4	AK_SCHPO	O60163 schizosach
4	491.5	17.5	AK_YEAST	P10869 saccharomyc
5	477.5	17.0	AKH_ECOI	P00861 escherichia
6	475.5	16.9	AKH_HAETN	P44505 haemophilus
7	463.5	16.5	AKH1_SERMA	P27125 serratia ma
8	461.5	16.4	AKH2_MAIZE	P49080 zea mays (m
9	452.5	16.1	AKH1_MAIZE	P49080 zea mays (m
10	440	15.6	AKH_BUCAP	P08674 buchnera ap
11	418	14.9	AK_THERM	P97151 thermus the
12	416.5	14.8	AK_COREP	O81911 corynebacte
13	416.5	14.8	AKH_BUCAP	O81911 corynebacte
14	415.5	14.8	AK1_CORGL	P26512 corynebacte
15	415.5	14.8	AK1_MYCSM	P41403 mycobacteri
16	412.5	14.7	AK1_CORGL	P41398 corynebacte
17	412.5	14.7	AKH_DAUCA	P37142 daucus caro
18	409.5	14.6	AK_MYCTU	P97048 mycobacteri
19	408.5	14.5	AK3_BACSU	P94417 bacillus su
20	402	14.3	AK2_BACST	O59229 bacillus sp
21	396	14.1	AK2_BACST	P53553 bacillus st
22	394	14.0	AKH_BUCAI	P57290 buchnera ap
23	380	13.5	AK1_BACSU	O04795 bacillus su
24	374	13.3	AK1_BSEAE	O06977 pseudomonas
25	368.5	13.1	AK2_BACSU	P08495 bacillus su
26	364.5	13.0	AK1_RICPR	O92217 rickettsia
27	347	12.3	AK2H_ECOI	P00562 escherichia
28	344	12.2	AK1_CHLTR	O84267 chlamydia t
29	340.5	12.1	AK1_AQUAR	O67221 aquifex aeo
30	335	11.9	AK1_HELPY	O92127 helicobacte
31	335	11.9	AK1_HELPY	O92127 helicobacte
32	334.5	11.9	AK1_CHLPM	O92110 chlamydia p
33	301	10.7	AK1_CHLMO	O9p332 chlamydia m

34	139	4.9	231	1	PYRH_THEMA	Q9X110 thermotoga
35	120.5	4.3	1280	1	MDRI_LEIRN	O06034 leishmania
36	113.5	4.0	753	1	YEHU_ECOI	P75764 escherichia
37	110.5	3.9	661	1	MCPI_BACSU	P39214 bacillus su
38	108.5	3.9	608	1	DXS_THEMA	Q9X291 thermotoga
39	108	3.8	378	1	GRK_HAEIN	P44507 haemophilus
40	105	3.7	227	1	PYRH_SULSO	Q97262 sulfolobus
41	105	3.7	462	1	MUTL_CLOTT	O05491 clostridium
42	105	3.7	867	1	DPOL_RICPR	O05949 rickettsia
43	104.5	3.7	381	1	GRK2_ECOI	P23524 escherichia
44	104.5	3.7	468	1	SG3_FOMAN	O6xk2d homo sapien
45	104.5	3.7	1853	1	CIPA_CLOTM	O06851 clostridium

ALIGNMENTS

RESULT 1	AK_METUA	STANDARD;	PRT;	473 AA.
AC	O57991;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Probable aspartokinase (EC 2.7.2.4) (Aspartate kinase).			
GN	M00571.			
OS	Methanococcus jannaschii.			
CC	Archaea; Euryarchaeota; Methanococci; Methanococcales;			
CC	Methanococcidococcaceae; Methanococcidococcus.			
CC	NCBI_TaxID=2190;			
CC	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=JAL-1 / DSM 2661 / ATCC 43067;			
RX	MDLINE=96337999; PubMed=8688087;			
RA	Bult C.J., White O., Olsen G.J., Zhou J., Fleischmann R.D.,			
RA	Sutton G.G., Blake O., Olsen G.J., FitzGerald L.M., Clayton R.A., Gocayne J.D.,			
RA	Scott J.L., Geoghagen N.S.M., Weidman J.F., Merrick J.M., Nguyen D.,			
RA	Overbeek R., Kirkness E.F., Tomb J.F., Adams M.D., Reich C.I.,			
RA	Utterback T.R., Kellley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,			
RA	Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,			
RA	Klenk H.-P., Fraser C.M., Smith H.O., Weiss C.R., Venter J.C.,			
RT	"Complete genome sequence of the methanogenic archaeon, Methanococcus			
RT	jannaschii";			
RL	Science 273:1058-1073(1996).			
CC	-1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-			
CC	aspartate.			
CC	-1- PATHWAY: First step in the common biosynthetic pathway leading			
CC	from Asp to diaminopimelate and Lys, to Met, to Ile and to Thr.			
CC	-1- SIMILARITY: Belongs to the aspartokinase family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U67506; AAB96565.1; -			
DR	PIR; C64371; C64371.			
DR	TIGR; M00571; -			
DR	InterPro; IPR001048; Aa_kinase.			
DR	InterPro; IPR002912; ACT.			
DR	InterPro; IPR005260; Asp_kin_monofn.			
DR	InterPro; IPR001341; Aspartate_kinase.			
DR	Pfam; PF00696; aak_kinase; 1.			
DR	Pfam; PF01842; ACT; 2.			
DR	TIGRPFAM; TIGR00656; asp_kin_monofn; 1.			
DR	TIGRPFAM; TIGR00657; asp_kinases; 1.			
DR	PROSITE; PS00324; ASPARTOKINASE; 1.			
KW	Transferase; Kinase; Threonine biosynthesis; Complete proteome.			
SQ	SEQUENCE 473 AA, 51392 MW, 811660FA5666C5F CRC64;			

Query Match 23.5%; Score 662; DB 1; Length 473;
 Best Local Similarity 32.3%; Pred. No. 8.3e-39;
 Matches 160; Conservative 110; Mismatches 167; Indels 58; Gaps 11;

QY 83 LSVWKKFGSSSVSSAARAEVAGLLTTPERRP--VYVLSAMGKTTNNLLAGEKAYCGC 140
 DB 1 MTTVMKFGSTVSGSERRHAKIVTKRKEDDDVVVVSAMSEVTHALVEISQAL--D 58
 QY 141 VTHVSEIEEM-NMVKSLIKTVDEL-----GLPIXCNTSLYEELQKGLAMKELT 191
 DB 59 VNDIAKVGDFIKFIEKHAKYKAEIKSEIEYKTIIDRIEIELEKVLIGVYLGELT 118
 QY 192 PRTSDYLVFGECKMSTRISAVLNKIRKARYDAFDIGFTTDEFGNA-----DILEAT 246
 DB 119 PXSRYILSFGERLSPISGAIKRDGKSLMEGEGAGITTDNNFSAKYLEVERL 178
 QY 247 YPAVAKRLHGDWITDPAIPVVTGFLGKWKSGAVTTLGRGSDLTATTIGKALGLREIQV 306
 DB 179 LPL-----LKEGIIIVVTGFIGT--TEEGYITTLGRGSDYSALIGYGLADIIIEI 228
 QY 307 WQDVGVLTCDPNITPHAKTVYLTFEERTELAYGAQVLPQSMRPARREGDIVRYKNS 366
 DB 229 WMDVSGVYTTDPLVATARIKPLSTTEAMELAYGARVLHPRITTEPMERGIPTLVKNT 288
 QY 367 YNPKAPGTLITRQDMXGLVVLTSIVLSKNTMLDIVSTMLGQYGLARVSGICYIED 426
 DB 289 FEPESEGTITNDMEMSDSIVAKIST--KNVALINIGAGVVGSTPAARIFKALGEE 346
 QY 427 LCISVDCVATSEVSUSLDPKINSRELIQASLDRVVELEK----- 471
 DB 347 VNVILISQSSSEFINIS-----LVSEEDVDKALKKKEFGDFPKKSFNNLT 394
 QY 472 IATVRLQORAITSLIG-NVEQSILTEKTRGKSGVNVOMISQASKYNNMLIYVDS 530
 DB 395 IRVSVYDKVQCVISVAGAKMRGAKIAGITFAVESGAGNIMIQSSSEVNTISRVIDEK 454
 QY 531 DAKALVALHQAFFE 545
 DB 455 DLINCVKLEHEKIE 469

RESULT 2
 AK3_ECOLI STANDARD; PRT; 449 AA.
 AC P08660;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lysine-insensitive aspartokinase III (Ec 2.7.2.4) (Aspartate kinase III).
 GN LYSC OR APK OR B4024.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RA MEDLINE=8611734; PubMed=3003049; Patte J.-C.;
 RA Cassan M., Patroc C., Cohen G.N., Patte J.-C.;
 RT "Nucleotide sequence of lysc gene encoding the lysine-sensitive aspartokinase III of Escherichia coli K12. Evolutionary pathway leading to three isofunctional enzymes";
 RT J. Biol. Chem. 261:1052-1057(1986).
 RU [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=K12 / MG1655;
 RA MEDLINE=94089392; PubMed=8265357;
 RA Blatter F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
 RA Daniels D.L.;
 RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the region from 89.2 to 92.8 minutes";

RL Nucleic Acids Res. 21:5408-5417(1993).
 RN [3]
 RP SEQUENCE OF 1-20 FROM N.A.
 RA MEDLINE=84015162; PubMed=6312411;
 RX Cassan M., Roncerea J., Patte J.-C.;
 RT "Nucleotide sequence of the promoter region of the E. coli lysc gene";
 RT Nucleic Acids Res. 11:6157-6166(1983).
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-aspartate
 CC -1- ENZYME REGULATION: SYNTHESIS AND ACTIVITY ARE SENSITIVE TO LYSINE, WHICH IS ONE OF THE END METABOLITES OF THE ASPARTIC ACID FAMILY
 CC BRANCHED PATHWAY.
 CC -1- PATHWAY: Lysine biosynthesis; first step.
 CC -1- SUBUNIT: Homodimer.
 CC -1- MISCELLANEOUS: ASPARTOKINASES I AND II ALSO CATALYZE THE SAME REACTION(S).
 CC -1- SIMILARITY: belongs to the aspartokinase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; M1812; AAA24095.1; -
 CC EMBL; U00006; AAC43118.1; -
 CC EMBL; AE000475; AAC76994.1; -
 CC EMBL; X00008; CAA24910.1; ALT_SEQ.
 CC PIR; G65209; KIECD3.
 CC Ecogene; EGI0550; lysc.
 CC InterPro; IPR001048; Aa_kinase.
 CC InterPro; IPR002912; ACT.
 CC InterPro; IPR005260; Asp_kin_monofn.
 CC InterPro; IPR001341; Aspartate_kinase.
 CC Pfam; PF00696; aakkinase; 1.
 CC Pfam; PF01842; ACT; 1.
 CC TIGRPFAM; TIGR00656; asp_kin_monofn; 1.
 CC TIGRPFAM; TIGR00657; asp_kinases; 1.
 CC PROSITE; PS00324; ASPARTOKINASE; 1.
 CC TRANSFERASE; Kinase; Lysine biosynthesis; Complete proteome.
 CC DOMAIN 1 245 ASPARTOKINASE III.
 CC FT DOMAIN 246 449 INTERPACE.
 CC FT CONFLICT 58 58 G->C (IN REF. 1).
 CC FT CONFLICT 401 401 G->A (IN REF. 1).
 CC FT SEQUENCE 449 AA; 48531 MW; 5B41CB3AE64D9848 CRC64;
 Query Match 20.2%; Score 569; DB 1; Length 449;
 Best Local Similarity 34.0%; Pred. No. 2.4e-32;
 Matches 161; Conservative 89; Mismatches 182; Indels 42; Gaps 14;

QY 85 VNMKFGSSSVSSAARAEVAGLLTTPERRP--VYVLSAMGKTTNNLLAGEKAYCGC 144
 DB 5 VVSKFGSTVADFDAMNSADIVSDANVR-LVVLASAGITNLVLAIE-----GLPEG 58
 QY 145 SEIEEMNMVKSILIKTVDELGLPIXCNTSLYEELQKGLAMKE-----LTPRTSDYLV 199
 DB 59 ERPEKIDARINQFALLERLRP---NVIREIERLLENITVLAEMAAALATSPALTDDELV 115
 QY 200 SPECKMSTRISAVLNKIRKARYQVADFDIGFTTDEFGND-----ILENTYAAVAKRL 254
 DB 116 SHEELMSTLLFEILIREDDVQAQWPDVAKV-WRTNDRGRAREPDIAALAEIALQILPRL 174
 QY 255 HGDWIDPAIPVVTGFLGKWKSGAVTTLGRGSDLTATTIGKALGLREIQWQDVGV 314
 DB 175 -----NEGIVITQGFISENK-GRTTLGRGSDYTAALAEALHSRVDIWDVGIY 227
 QY 315 TCDPNITPHAKTVYLTFEERTELAYGAQVLPQSMRPARREGDIPRYKNSYKAPGT 374
 DB 228 TTDPRVVSAAKRIIDIAFAEAEMKATFAKVLHPATILLPAVRSDIPFVVGSSKDPKAGGT 287

QY 375 LITRQDMDXGLVLTSLVSKSNMTMDIVSRMIGQYGLARVSGICYEDLCISYDCV 434
 DB 288 LVCNTE---NPPFLPALAKRNQTLTLTSLNLMHSGFLAEVFGILARN--ISVDLI 342
 QY 435 ATSEVSVSLDPSKIMSR--ELIQASELDHVEELFKIAIVRLQORALISLIGN-VE 491
 DB 343 TISEVSVALTLDSTSTGDTLLNQS-----LMELSLALCRVEVEEGALVALIGNDLS 397
 QY 492 QGSLILEKTRVLRSGVNVQVMSIGASKVMSLIVHSDKALVEALHQAFFE 545
 DB 398 KAGVKEVFGVLER--FNIRMICYGASSHNCFLVPSEDAQVQVQLHSLNLF 449
 RESULT 3
 AK_SCHPO STANDARD; PRT: 519 AA.
 ID AK_SCHPO
 AC 060163;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable aspartokinase (EC 2.7.2.4) (Aspartate kinase).
 GN SPBC19F5.04
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxId=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=218401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras P., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., Meleis J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Shelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymporter B.,
 RA Weidens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Fumelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Koehet M., Galliard C., Talada V.A., Garçon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez X., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Sipakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RA "the genome sequence of Schizosaccharomyces pombe";
 RA Nature 415:871-880(2002).
 RL Nature 415:871-880(2002).
 CC -1- CATALYTIC ACTIVITY: ADP + L-aspartate = ADP + 4-phospho-L-
 CC aspartate.
 CC -1- PATHWAY: First step in the common biosynthetic pathway leading
 CC from Asp to diaminopimelate, to Met, to Ile and to Thr.
 CC -1- SIMILARITY: Belongs to the aspartokinase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL databank on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, AL022599; CAA18652.1; -
 CC DR PIR, T39822; T39822.

DR GenedB, Spombe, SPBC19F5.04; -
 DR InterPro: IPR001048; Aa_kinase.
 DR InterPro: IPR002912; ACT.
 DR InterPro: IPR001341; Aspartate_kinase.
 DR Pfam: PF00696; aak_kinase; 1.
 DR Pfam: PF01842; ACT; 2.
 DR TIGRfam: TIGR00657; asp_kinase; 1.
 DR PROSITE: PS00324; ASPARTOKINASE; 1.
 KW Transferase; Kinase; Threonine biosynthesis.
 SC SEQUENCE 519 AA; 56812 MW; 0B392EB4A8BCFCB CRC64;
 Query Match 18.4%; Score 516.5; DB 1; Length 519;
 Best Local Similarity 28.9%; Pred. 1.4e-28;
 Matches 152; Conservative 107; Mismatches 186; Indels 81; Gaps 20;
 QY 85 VVMKFGSSVSS-AARMA-EVAGLILTFPEERFVVLAMGK-----TNNLLAGEKA 136
 DB 16 VVQKFGTSGVKFP-KIAVDVAKYLS--TKRVALVCSASTDTRKAGTTRILRTEAA 73
 QY 137 VCCGVINSEIENMMVMSLHKT---VDLQ---LPXICNTSLYEQLKGIAMK 188
 DB 74 LRPVAGSVHDLV--RIETDVQARDFIQDVGIQDELIDAFHADCVELDQYLAIRVLS 131
 QY 189 ELTPRTSDVLSFGECWSTRIFSAYLNKIRVKAQYDAFDIGFTTD--EFGNADI--- 242
 DB 132 EVSPRTDLVIGMGRSLSCRFMAVLKDQIGDS---EFIDMSHIDQREWRMLDSFYA 188
 QY 243 ----LEATYPAVARLHGMDIODPAIPVVTGFGKMGSGAVTTLGRGSDLTATTIGK 297
 DB 189 YLASGLASKVTVAGNK-----VVVVGFGFM--VPGGLSQIGRGYTDPCALLAV 237
 QY 298 ALGLREIOWKVDQVGLTCDPNIPYHAKTVYTLFEEBATELAVGAQVLHPOSRRPAREG 357
 DB 238 GLNABEQIWKVEVDIFPADRKPTALRLTLFPEBAELTYGSEVHPFTMSGVHA 297
 QY 358 DIPVRKVSYNKAPGTLI---TROR-----DMXGLV--VLISIVLKS 397
 DB 298 RIPRIKVNQNRGKGTIVFPDITSRHSGATPPHPKIMPDIDIASLANKGATAVITKOT 357
 QY 398 VMLDVSVRMLGQVGLFARVSGICYEDLCISYDCVATSEVSVSLDPSKIMSRBLIQ 457
 DB 358 IMVINQSRKISAHGFLASIPAL--LKYKLADLITTSVHVSALYEE-----S 407
 QY 458 QASELDHVEELEKIAIVRLQORALISLIGNVEGSSILILEKTRV---LRKSGVNVQMI 514
 DB 408 DQGNHAEARVELRRLGTIDILHGLATLSLVGKHEMT--TSVAGRMFCKLAEAQINTEMI 465
 QY 515 SQGASKVMNSLIVHSDKALVEALHQAFFEDVLSQVEAE-NILIV 559
 DB 466 SQGASEINISCVIDEKMAVKALNVHKELEPLALHEVPSQSMILV 511
 RESULT 4
 AK_YEAST STANDARD; PRT: 527 AA.
 ID AK_YEAST
 AC P10869;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Aspartokinase (EC 2.7.2.4) (Aspartate kinase).
 GN HCM3 OR YER052C
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxId=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88115350; PubMed=2892836;
 RA Rafalski J.A., Falco S.C.;
 RA "Structure of the yeast HCM3 gene which encodes aspartokinase";
 RA J. Biol. Chem. 263:2146-2151(1988).
 RL [2]
 RN REVISIONS.

RX MEDLINE=90368723; PubMed=2168408;
 RA Rafalski J.A., Falco S.C.;
 RT "Structure of the yeast HOM3 gene which encodes aspartokinase.";
 RL J. Biol. Chem. 265:15346-15346(1990).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=6286c / AB972;
 RX MEDLINE=97313264; PubMed=9169868;
 RA Dietrich F.S., Mulligan J.T., Henessey K.M., Yelton M.A., Allen E.,
 Araujo R., Aviles B., Benno A., Brennan T., Carpenter J., Chen E.,
 Chery J.M., Chung E., Duncan M., Guman E., Hartzell G.,
 Hummel-Smith S., Hyman R.W., Kayser A., Kemp C., Leachart D., Lew H.,
 Lin D., Mosedale D., Nakahara K., Natch A., Norgren R., Oetner P.,
 Oh C., Petel F.X., Roberts D., Sehl P., Schamm S., Shogren T.,
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";
 RL Nature 387:78-81(1997).
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
 aspartate.
 CC -1- PATHWAY: First step in the common biosynthetic pathway leading
 CC from Asp to diaminopimelate, to Met, to Ile and to Thr.
 CC -1- SIMILARITY: Belongs to the aspartokinase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to a license@sib.ch).
 CC
 DR EMBL: J03526; AAA34681.1; -
 DR EMBL: U18796; AAB64587.1; -
 DR PIR: A35888; KTBXD.
 DR GerMOnline: 139132; -
 DR SGD: S0000854; HOM3.
 DR InterPro: IPR001048; Aa_kinase.
 DR InterPro: IPR002912; ACT.
 DR InterPro: IPR001341; Aspartate_kinase.
 DR Pfam: PF00696; aak_kinase; 1.
 DR Pfam: PF01842; ACT; 2.
 DR TrEMBL: TIGR00657; asp_kinase; 1.
 DR TrEMBL: PS00324; ASPARTOKINASE; 1.
 DR TrEMBL: KINASE; Threonine biosynthesis.
 KM Transferase; Kinase; Threonine biosynthesis.
 SQ SEQUENCE 527 AA; 58109 MW; D4D28F8D4374898 CRC64;

Query Match 17.5%; Score 491.5; DB 1; Length 527;
 Best Local Similarity 28.1%; Pred. No. 7, 7e-27;
 Matches 143; Conservative 107; Mismatches 180; Indels 79; Gaps 14;

QY 85 VNMKFGSSVSSAARAEVAGLILF-----PEERPVVYLSMKG-----TTNNL--- 130
 DB 15 VVQKFGTSTYKGFPP-VQIVDDIVKHYSKPDGPNVNAVCSARSYTAEGTSTLLKCC 73
 QY 131 -LAGEAVAGGVV-----HVSEIEEMNVKSLHKTVDGLKFLPCTSTLYEAEOLKKG 183
 DB 74 DLASQSEFODILEVRQDHIDNADFLINPLQKLVND-----TKKELELYKYLINA 127
 QY 184 IANMKELPTSDIYVSPCEMSTRIFSAYLKIRKARQYDAFIFGTTTDEFGNADIL 243
 DB 128 SKVLGEVSRITVDLVMSCKEKLSCLEMTALCDROGKAKY--VDLSHIVSDPFGASALD 184
 QY 244 EATYPAVAKELHGDW-----IODPAIVVTGFLGKMGKSGAVTTTGRGSDPLTATTIKA 298
 DB 185 NSRTYTLVQALKLKLAPVSAKRIVPVGTGFFGL-VPTGLINGRGRTDUCALINVA 243
 QY 299 LGRREIQWCDVDVGLTCDPNIYPAKIVPYLTFEBATELAYFGAQLVHPSGMRPAREGD 358
 DB 244 VNDELQWKEVDGIFLADPRKVPPEARLIDSVTPBASLTYGSEVILHPTMEQVIRAK 303
 QY 359 IPRVAKSYNPKAPGLI-----TRQDMXGLVIVLSI 392
 DB 304 IPIRKVQNPGLNGITITIPDNVAKKGESTPPHPENNLSSESYERKKG-----ATAI 356

QY 333 VLKSNVTMDIVSTRMLGQYGFLLARVSGICVIEDLCISDVCVATSEVSVSLDPKIMS 452
 DB 357 TTKNDIFVINIHNSKTKLSHGFLAQIFTL--LQRYKLVLDISTSEVHSMALPIPDADS 414
 QY 453 RELIQASELDHVEEELKIAIVLLOQRAIISLIG-NVQSLSILEKTRGVLRKSGVNV 511
 DB 415 LKSLRQAE-----EKRIILGSVDITKSLSTVVGKMKQYIGIAGTMTTLAEGINI 468
 QY 512 QMISGASKVMSLIVHSDQAKALVEALH 540
 DB 469 EMISGANEINISCVINESDSIKALQCIH 497

RESULT 5
 AKIH_ECOLI STANDARD; PRT; 820 AA.
 ID AKIH_ECOLI
 AC P00561; Q47659;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Bifunctional aspartokinase/homoserine dehydrogenase I (AKI-HDI)
 DE [includes: Aspartokinase I (EC 2.7.2.4); Homoserine dehydrogenase I
 DE (EC 1.1.1.3)].
 OS THRA OR THRA1 OR THRA2 OR B0002.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_Taxid=562;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8107247; PubMed=7003595;
 RA Katioka M., Cossart P., Sibilli L., Saint-Girons I., Chalvignac M.A.,
 RA le Bras G., Cohen G.N., Yaniv M.;
 RT "Nucleotide sequence of the thra gene of Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 77:5730-5733(1980).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=92334977; PubMed=1630901;
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
 RA Isono K., Mizobuchi K., Nakata A.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of
 RT the 0-2.4 min region.";
 RL Nucleic Acids Res. 20:3505-3508(1992).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=95334362; PubMed=7610040;
 RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
 RA Blattner F.R.;
 RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
 RT region from 92.8 through 100 minutes.";
 RL Nucleic Acids Res. 23:2105-2119(1995).
 RN (4)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80077291; PubMed=390305;
 RA Cossart P., Katioka M., Yaniv M.;
 RT "Construction and expression of a hybrid plasmid containing the
 RT Escherichia coli thra and thir genes.";
 RL Mol. Gen. Genet. 175:39-44(1979).
 RN (5)
 RP SEQUENCE OF 51-129.
 RX MEDLINE=80043179; PubMed=387092;
 RA Sibilli L., le Bras G., Cossart P., Chalvignac M.A., le Bras G.,
 RA Briley P.A., Cohen G.N.;
 RT "The primary structure of Escherichia coli K 12 aspartokinase I -
 RT homoserine dehydrogenase I: sequence of cyanogen bromide peptide CB
 RT 3.";
 RL Biochimie 61:733-739(1979).
 RN [6]
 RP REVISION TO 11.
 RX MEDLINE=83135751; PubMed=6298218;

RA Cossart P.;
 RL Unpublished results, cited by:
 RL Zakin M.M., Duchange N., Ferrara G.N.;
 RL J. Biol. Chem. 258:3028-3031(1983).
 CC -1- CATALYTIC ACTIVITY: L-homoserine + NAD(P) (+) = L-aspartate 4-
 CC semialdehyde + NAD(P)H.
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
 CC aspartate.
 CC -1- ENZYME REGULATION: The enzyme activities are regulated
 CC allosterically by L-threonine.
 CC -1- PATHWAY: Catalyzes 2 nonconsecutive reactions in the common
 CC biosynthetic pathway leading from Asp to diaminopimelate and Lys,
 CC to Met, and to Thr and Ile.
 CC -1- SUBUNIT: Homotetramer.
 CC -1- MISCELLANEOUS: Aspartokinase II-homoserine dehydrogenase II and
 CC aspartokinase III also catalyze the same reaction(s).
 CC -1- SIMILARITY: In the N-terminal section, belongs to the
 CC aspartokinase family.
 CC -1- SIMILARITY: In the C-terminal section, belongs to the
 CC homoserine dehydrogenase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: V00361; CAA23660.1; ALT_SBO.
 DR EMBL: X68872; CAA48734.1; -
 DR EMBL: D10483; BAB96579.1; -
 DR EMBL: U14003; AAA97301.1; -
 DR EMBL: AE000111; AAC73113.1; -
 DR EMBL: V00360; CAA23659.1; -
 DR EMBL: M28570; AAA24673.1; -
 DR EMBL: M10644; AAA24671.1; -
 DR PIR: B64720; DEECK.
 DR EcoGene: EG10998; thra.
 DR InterPro: IPR001048; Asp_kinase.
 DR InterPro: IPR002912; Act.
 DR InterPro: IPR001341; Aspartate_kinase.
 DR InterPro: IPR001342; Homoserine_dh.
 DR InterPro: IPR005106; NAD_Binding_3.
 DR Pfam: PF00696; aak_kinase_1.
 DR Pfam: PF01842; ACT_1.
 DR Pfam: PF00742; Homoserine_dh_1.
 DR Pfam: PF03447; NAD_Binding_3_1.
 DR TIGRfams: TIGR00657; asp_kinases_1.
 DR PROSITE: PS00324; ASPARTOKINASE; 1.
 DR PROSITE: PS01042; HOMOSER_DHGEMASE; 1.
 KW Transferase; Kinase; Oxidoreductase; Threonine biosynthesis; NADP;
 KW Allosteric enzyme; Multifunctional enzyme; Complete proteome.
 FT DOMAIN 1 249 ASPARTOKINASE I.
 FT DOMAIN 250 470 INTERPASE.
 FT DOMAIN 471 820 HOMOSERINE DEHYDROGENASE I.
 FT NP_BIND 471 478 NADP (POTENTIAL).
 FT CONFLICT 113 113 Q -> E (IN REF. 5).
 FT CONFLICT 230 230 D -> N (IN REF. 1, 2 AND 4).
 FT CONFLICT 375 375 Q -> L (IN REF. 1, 2 AND 4).
 FT CONFLICT 393 393 T -> A (IN REF. 1, 2 AND 4).
 FT CONFLICT 406 406 M -> A (IN REF. 1, 2 AND 4).
 FT CONFLICT 553 553 D -> N (IN REF. 1 AND 2).
 FT CONFLICT 588 588 DY -> IT (IN REF. 4).
 FT CONFLICT 607 607 TY -> I (IN REF. 1, 2 AND 4).
 FT CONFLICT 607 607 T -> R (IN REF. 1, 2 AND 4).
 FT CONFLICT 658 658
 FT CONFLICT 820 AA; 89119 MW; 08F28B9ECCAB10ED CRC64;
 SQ SEQUENCE

Query Match 17.0%; Score 477.5; DB 1; Length 820;
 Best Local Similarity 29.4%; Pred. No. 1.3e-25;
 Matches 142; Conservative 98; Mismatches 210; Indels 33; Gaps 13;
 86 VMKFGSSVSSAPAAVAGLILPPEERPV-VVLSAMGKTTNNLLAGEKAVG--CGVI 142

DB 3 VAKFGSTVSNRERLFRVADLIESNARGQVATVLSAPAKITNHLVMIEKTIISGQDALP 62
 143 HUSEIEEMWVKSLHIKTYDEIGLP-----XICNTSYLEQLKJGAMKELPRRSYD 197
 DB 63 NISDAER-IFAEILTGLAAAGPFPALQKTFVQEPFAQIGVHLSLGGCCDSINAA 121
 198 LVSFPGCWSTRIFSAVINKIRVAKQYDAFDIG---FITDEF--GNADILEATYPAVA 251
 DB 122 LIRGKMSIALMAGL-----EARGHVTVIDPEKLLAVGHYLESTVDIAESTRIIA 176
 252 KRLHGMIDDPALPVVTGFLGKWKSGAVTTVGRGSGSLTTTGKALGLEIOWKMDVD 311
 DB 177 SRIPAD-----HVMVMAAGFTA-GNEKGELIVTGRNGSYSAVAACIRADCCIMDVD 230
 312 GVTCTDPNTPYPAKVTVPYTFEEAELEIVFGQVHPOSPAPAREGDIPIVKNVSKPA 371
 DB 231 GVTCTDPNTPYPAKVTVPYTFEEAELEIVFGQVHPOSPAPAREGDIPIVKNVSKPA 290
 372 PGLTITRORDMXGAVLTSTIVLKSNTVMTLDIVSTRMLGQYGLARVSGICYEDLCISV 431
 DB 291 PGLTIGASRDEDE-LPVKGISNLNNMAFVSVCGBMGVGMARVAMSRARISVYL 348
 432 DCVATSEVSVSVSLDPSK-INSRELIQASLELDHVEELEKTAIVRLQRAITSLGN- 489
 DB 349 ITGSSEYSISFCVPQSDCVRAERAMQEEFYELKEGLLEPLAVERL--AITSVGDG 405
 490 VEGSSLTLEKTRGVLRKSGVNTVMTSQGASKYMSLIVHSDAKALVALHQAFFEDVL 549
 DB 406 MRLRIGSAKFFAALARAANINIVAIAGSSSEISIVVNNDDATGAVTHQMLFNTDQV 465
 550 SQV 552
 DB 466 IEV 468

RESULT 6
 AKH_HAEIN STANDARD; PRT; 815 AA.
 AC P44505;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Bifunctional aspartokinase/homoserine dehydrogenase (AK-HD)
 DE [includes: Aspartokinase (EC 2.7.2.4); Homoserine dehydrogenase
 DE (EC 1.1.1.3)].
 GN THRA OR H10089.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; Pubmed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utecherback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RL Science 269:496-512(1995).
 CC -1- CATALYTIC ACTIVITY: L-homoserine + NAD(P) (+) = L-aspartate 4-
 CC semialdehyde + NAD(P)H.
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
 CC aspartate.
 CC -1- PATHWAY: Catalyzes 2 nonconsecutive reactions in the common
 CC biosynthetic pathway leading from Asp to diaminopimelate and Lys,

CC to Met, and to Thr and Ile.
 CC -1- SUBUNIT, Homocitramer (By similarity).
 CC -1- SIMILARITY: In the N-terminal section; belongs to the
 CC aspartokinase family.
 CC -1- SIMILARITY: In the C-terminal section; belongs to the homoserine
 CC dehydrogenase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U32694; AAC21767.1; -
 DR PIR; A64048; A64048.
 DR TIGR; H10089; -
 DR InterPro; IPR001048; As. kinase.
 DR InterPro; IPR002912; ACT.
 DR InterPro; IPR001344; Aspartate_kinase.
 DR InterPro; IPR001342; Homoserine_dh.
 DR InterPro; IPR005106; NAD_binding_3.
 DR Pfam; PF00696; aakkinase; 1.
 DR Pfam; PF01842; ACT; 2.
 DR Pfam; PF00742; Homoserine dh; 1.
 DR Pfam; PF03447; NAD binding 3; 1.
 DR TIGRfam; TIGR00657; asp_kinase; 1.
 DR PROSITE; PS00324; ASPARTOKINASE; 1.
 DR PROSITE; PS01042; HOMOSER-DHGENASE; 1.
 DR Transferrase; Kinase; Oxidoreductase; Threonine biosynthesis; NADP;
 KM Allosteric enzyme; Multifunctional enzyme; Complete proteome.
 FT DOMAIN 1 249 ASPARTOKINASE.
 FT DOMAIN 250 470 INTERPACE.
 FT DOMAIN 471 815 HOMOSERINE DEHYDROGENASE.
 FT NP_BIND 471 478 NADP (POTENTIAL).
 FT SEQUENCE 815 AA; 88220 MW; 7FD5512CC4BFF7 CR664;
 SQ
 Query Match 16.9%; Score 475.5; DB 1; Length 815;
 Best Local Similarity 29.1%; Pred. No. 1.8e-25;
 Matches 142; Conservative 103; Mismatches 188; Indels 55; Gaps 16;
 QY 86 VMKRGSSVSSAARMAEVALI-LTPPEEPVYVLSAMKTNINLLAGEKAV-GCGVI 142
 DB 3 VLKFGGSLANPERFSQAALIEQAHLEBPAGVLSAPAKNTNHLVALSEKALNQSTDT 62
 QY 143 HVSEIEE--NNWVKS.LHIKT--VDEIGLPIYICNTSLYEELQILKGIAMKELTPTSDYL 198
 DB 63 HENNAIEIFNYINIGLHTENNQPDINGTKALIDAEFVQIKGLEIRQAKVEDAVKATI 122
 QY 199 VSPGECMSTRIFSAVLYNKIKVAKQYDAPI-----GTTTDEGQNDILEATYP 248
 DB 123 DCRBEKLSIAMKAMF-----EARGSVHIVDPVKOLAKGGLSE--SVEIEEST-- 171
 QY 249 AVAKRLHGMVIOPEAIPVYTGFLGKWKSGAVTTLGRGSDLTATTIGKALGREIQWK 308
 DB 172 ---KRVDAANIARKKVLMAFGTR-GNEKEVLJLRNSDYSAAALACLASACEIWT 227
 QY 309 DVDGVLTCDEPIYHAKTVPLTTEATELAFGAQVYLHPQSKRPAREGDIYRVKNSIN 368
 DB 228 DVDGVYTCDEPLVADALPLTLSTREAMELSYFGARVIRRTIGPLLPNITCVIKNTGN 287
 QY 369 PKARGLTITRODMDXGLVYLTSVLR-----SNVTMLDIVSTRMLGQVGLFARVSGICY 423
 DB 288 PSAPGSIT-----DENVASESIQVGINLNDLMAFNVSQGMQGMASRVPSAAS 340
 QY 424 IEDICISVDCVATEVSVSLDSKISWPELLIQASSELDRHYE---ELEKIAIYELIQ 479
 DB 341 GAGISVILITQSSSEYSSISFCV-FVK--SAEVAKTVLETFANELNEHOLEPELVKDL- 396
 QY 480 QRAITSLIGN-VEOSSLILEKTVLRKSGVNTQMSOGASKVNMSLIYHSDAKALVYA 538
 DB 397 --SIISVGDGMKQAKGIAAFPSALMQANISITVALAQSSSESSISAVVPQMKALREAVKA 454

QY 539 LHOAFPED 546
 DB 455 THQALENN 462
 RESULT 7
 AKI_H_SERMA STANDARD. PRT. 819 AA.
 ID AKI_H_SERMA STANDARD. PRT. 819 AA.
 AC F27725; Q59936; Q59937; Q60127; Q60157;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Bifunctional aspartokinase/homoserine dehydrogenase I (AKI-HDI)
 DE [Includes: Aspartokinase I (EC 2.7.2.4); Homoserine dehydrogenase I
 DE (EC 1.1.1.3)]
 GN THRA OR THRAL OR THRA2.
 OS Serratia marcescens.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Serratia.
 CX NCBI_TaxId=615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=841;
 RX MEDLINE=9319048; PubMed=8423151;
 RA Omori K., Suzuki S., Komatsubara S.;
 RT "Nucleotide sequence of the Serratia marcescens threonine operon and
 RT analysis of the threonine operon mutations which alter feedback
 RT inhibition of both aspartokinase I and homoserine dehydrogenase I.";
 RL J. Bacteriol. 175:785-794(1993).
 CC -1- CATALYTIC ACTIVITY: L-homoserine + NAD(P)(+) = L-aspartate 4-
 CC semialdehyde + NAD(P)H.
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
 CC aspartate.
 CC -1- ENZYME REGULATION: The enzyme activities are regulated
 CC allosterically by L-threonine.
 CC -1- PATHWAY: Catalyzes 2 nonconsecutive reactions in the common
 CC biosynthetic pathway leading from Asp to diaminopimelate and Lys,
 CC to Met, and to Thr and Ile.
 CC -1- SUBUNIT: Homocitramer.
 CC -1- SIMILARITY: In the N-terminal section; belongs to the
 CC aspartokinase family.
 CC -1- SIMILARITY: In the C-terminal section; belongs to the homoserine
 CC dehydrogenase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D10385; BAA38474.1; -
 DR EMBL; D10386; BAA38477.1; -
 DR EMBL; D10387; BAA38480.1; -
 DR EMBL; X60821; CAA43212.1; -
 DR PIR; B47057; B47057.
 DR InterPro; IPR001048; As. kinase.
 DR InterPro; IPR002912; ACT.
 DR InterPro; IPR001344; Aspartate_kinase.
 DR InterPro; IPR001342; Homoserine_dh.
 DR InterPro; IPR005106; NAD_binding_3.
 DR Pfam; PF00696; aakkinase; 1.
 DR Pfam; PF01842; ACT; 2.
 DR Pfam; PF00742; Homoserine dh; 1.
 DR Pfam; PF03447; NAD binding 3; 1.
 DR TIGRfam; TIGR00657; asp_kinase; 1.
 DR PROSITE; PS00324; ASPARTOKINASE; 1.
 DR PROSITE; PS01042; HOMOSER-DHGENASE; 1.
 DR Transferrase; Kinase; Oxidoreductase; Threonine biosynthesis; NADP;
 KM Allosteric enzyme; Multifunctional enzyme.
 FT DOMAIN 1 249 ASPARTOKINASE I (BY SIMILARITY).
 FT

FT DOMAIN 250 470 INTERPACE (BY SIMILARITY).
 FT DOMAIN 471 819 HOMOSERINE DEHYDROGENASE 1 (BY
 FT SIMILARITY).
 FT NP BIND 471 478 NADP (POTENTIAL).
 FT VARIANT 330 330 G -> D (IN STRAIN HNR21; LOSS OF
 FT INHIBITION).
 FT VARIANT 352 352 S -> F (IN STRAIN TUR156; LOSS OF
 FT FEEDBACK INHIBITION).
 FT VARIANT 479 479 A -> T (IN STRAIN HNR59; THR-RESISTANT
 FT HDI).
 SQ SEQUENCE 819 AA; 88494 MW; 1F18552B036AE39 CRC64;
 Query Match 16.5%; Score 463.5; DB 1; Length 819;
 Best Local Similarity 27.5%; Pred. No. 1.3e-24;
 Matches 137; Conservative 111; Mismatches 186; Indels 65; Gaps 17;
 QY 86 VMKGGSSVSAARMVAGLITFPEERPV-VILSANGKTNINLLAGEKAV-GCGVI-142
 DB 3 VLFKGTSVANARFIRVADIMESNARQGVATVLSAPKINHLVAMIDKTVAGQDILP 62
 QY 143 HVSEIE-----ENNMYSLHKTVDLGLPXCNTSLYELBQLKGI 185
 DB 63 NMSDARIRPADLISLAQALPGFHYDLKGV-----VFQKQYLGHSV 109
 QY 186 MKMELPTPSDVIYVSGECMSTRIFAVYLNKIRVAKQYDARDIG---FITTEF--GN 239
 DB 110 LUGQCPDSVNAATTCGEKLSIAMEGV---PAKGYPTVIVPEKYLAKGHYLEST 164
 QY 240 ADILEATYPAVANKELHGDWIODPAIPVTVFLGKWKSGAVTLGRGSDLTATTIGAL 299
 DB 165 VDLAESTLRLAAAIAD-----HYLMAGFTA-GNDKGLVLLGRNSDYSAAVLAACL 218
 QY 300 GLREIQVMKDVGSLTCDPRITYPHATVYLFPEETELAYGAQVLPQSRPAREDI 359
 DB 219 RADCCETWDVDVYTCDPRTYDARLLKMSVQSELMETSYGAVALHRTTPTPAQPQI 278
 QY 360 PVRVKSYPKAPGTILTR-QRDMXGLVLTSLVSKSVNTLMDIVSTRMLQGYGLARV 418
 DB 279 PCLIKTNSNPQABGTILGKDSITADHPVGIINL---NNMAININSGGKMGVMARV 335
 QY 419 SGCIYBDCISVDCVATSEVSVSLVDSKIV-SRELIQASSEL---DHVELEIKAI 474
 DB 336 FAVMSRAGISVAVITQSSSEYSSISFCVQGEORARALKEEFYELKQGVLDPLD--- 391
 QY 475 VRLQORATISLIGN-VEQSSLTLEKTVLRKSGVNVOMISQASXVMMSLIYVDSAK 533
 DB 392 --VMERLAIISVGDGKRTIRGISARFPALAPAININIVAIQSSSESSISVYVDSAT 449
 QY 534 ALVEALHQAFEDVLSQV 552
 DB 450 TGVVSHQMLENTDQVLEV 468
 RESULT 8
 AKH2 MAIZE STANDARD; PRT; 917 AA.
 ID AC P49080;
 AC 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Bifunctional aspartokinase/homoserine dehydrogenase 2, chloroplast
 DE precursor (AK-HB 2) (AK-HSDH 2) [includes: Aspartokinase (EC 2.7.2.4);
 DE Homoserine dehydrogenase (EC 1.1.1.3)].
 GN AKHSDH2.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Paccad clade; Panicoideae; Andropogoneae; Zea.
 CX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N. A.
 RC TISSUE=Seedling leaf;
 MEDLINE=55148730; PubMed=7846152;

RA Muehlbauer G.J., Somers D.A., Matthews B.F., Gengenbach B.G.;
 RT "Molecular genetics of the maize (Zea mays L.) aspartate kinase-
 RL homoserine dehydrogenase gene family.";
 RL Plant Physiol. 106:1303-1312(1994).
 CC -1- CATALYTIC ACTIVITY: L-homoserine + NAD(P)(+) = L-aspartate 4-
 CC semialdehyde + NAD(P)H.
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
 CC aspartate.
 CC -1- PATHWAY: Catalyzes 2 nonconsecutive reactions in the common
 CC biosynthetic pathway leading from Asp to diaminopimelate and Lys,
 CC to Met, and to Thr and Ile.
 CC -1- SUBUNIT: Homo- or heterodimer (Potential).
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: In the N-terminal section; belongs to the
 CC aspartokinase family.
 CC -1- SIMILARITY: In the C-terminal section; belongs to the homoserine
 CC dehydrogenase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L33913; AAA74361.1; -
 CC PIR: T02954; T02954.
 DR MAIZEDB; 66609;
 DR InterPro: IPR001048; Aa_kinase.
 DR InterPro: IPR002912; ACT.
 DR InterPro: IPR001341; Aspartate_kinase.
 DR InterPro: IPR001342; Homoserine_dh.
 DR InterPro: IPR005106; NAD_binding_3.
 DR Pfam: PF00696; aakinese; 1.
 DR Pfam: PF01842; ACT; 2.
 DR Pfam: PF00742; Homoserine_dh; 1.
 DR Pfam: PF03447; NAD_binding_3; 1.
 DR TIGRPFAMs: TIGR00657; asp_kinases; 1.
 DR PROSITE: PS00324; ASPARTOKINASE; 1.
 DR PROSITE: PS01042; HOMOSER DEHYDASE; 1.
 KW Transferase; Kinase; Oxidoreductase; Methionine biosynthesis; NADP;
 KW Multigene family; Chloroplast; Transist peptide;
 KW Multigene family.
 FT TRANSIT 1 89 CHLOROPLAST (POTENTIAL).
 FT TRANSIT 90 917 BIFUNCTIONAL ASPARTOKINASE/HOMOSERINE
 FT TRANSIT 90 917 DEHYDROGENASE 2.
 FT TRANSIT 90 917 ASPARTOKINASE.
 FT TRANSIT 90 917 INTERPACE.
 FT TRANSIT 90 917 HOMOSERINE DEHYDROGENASE.
 FT TRANSIT 90 917 NADP (POTENTIAL).
 FT TRANSIT 90 917 NADP (POTENTIAL).
 SQ SEQUENCE 917 AA; 99582 MW; 69225EAL3CAE7E9 CRC64;
 Query Match 16.4%; Score 461.5; DB 1; Length 917;
 Best Local Similarity 26.3%; Pred. No. 2e-24;
 Matches 153; Conservative 114; Mismatches 222; Indels 93; Gaps 21;
 QY 18 IPPASSGHVRLGACFTGTGPRGANGLSMVVADTSRRA-----KQADGGGVGAPVLS 72
 DB 10 LPPAAA-----AAWRPR-----ASSNRAVAVIQCWYELSDHYLGGPLRI 51
 QY 73 GLMEG-----GOLSVVMKFGSSVSAARMAY 103
 DB 52 GQGGSLHHRSTNFIAPAAAIISVQDEVTNVLKQGMWS-VHFGGTCTGTPRRICV 110
 QY 104 AGILTFPEERPVVILSANGKTN--NLLAGEKAVGCGVHNSIEEMMYSLHIXT 160
 DB 111 ANIVLDSSEKRIIIVSAMSRYTDMVYNVQVQGRDSDSYAIALAIVEKMTAKDL 170
 QY 161 VDELG-LPKICNTSLVELEQLKGIAMKMLPTPSDVIYVSGECMSTRIFSAVYLNKIRV 219
 DB 171 GEDLARPLSGLHSDVSNLQAMLRALYIAGHATLESFDFVGHGILMSQMLSYAIKSGA 230

QY 220 KARQYDAPDIGITITDEFGNADILEATYPAVAKLHGDWITODPA-IPVITGFLGKWKSG 278
 DB 231 PCSMDTREVLVITSGGNQVD--PDYLECEKTLQKFKFSRQFALIVAGFIAS--TAG 285
 QY 279 AV-TTLGSGSLTITTTGKALGLREIQWVDVGLTODPNITPHAKTIPVLTPEEATE 337
 DB 286 NPTTLKRDGSDPSAIVGSLVPAQVITMTDVGVSADPRKSEAVITLSTISYQAME 345
 QY 338 LAFYGAQVLHPQSKPAREGDIPIRVKNSYNPKAPGTLITRQ----RDMXGLVLTST 392
 DB 346 MSYFGANVLPRTIIPVKMDNPIVIRMTFSLAPGTWICQPANENGDD--ACVKSF 402
 QY 393 VLKSNVTMDIVSTKYLQGYGLAVSGICIEDLCISYDVC--ATSEVSVSLDPSKI 450
 DB 403 ATVDNIALVNVGEGTGAAGVPGTASAI--PSAVKQVGNVIMISQASSEHSVCFAV-PEK- 458
 QY 451 MSRELIQOASEL-DHVEELE--KIAIVRLLOQRAIISLIQ-----NVEQSSLILEKTGR 502
 DB 459 ---EVAVVAEHLHDFREALAGRLSKVEVINGCSIIAAGLRMASTPGVSAIIFD---- 511
 QY 503 VLKSGVNTQMTSQASKVNSLIYHSDAKALVBAHQAF 544
 DB 512 ALAKANINRAIAGGSEYNTIVLKQDCVRAIIRAHSRFF 553
 RESULT 9
 AKH1_MAIZE
 ID AKH1_MAIZE STANDARD; PRT; 920 AA.
 AC P49079;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Bifunctional aspartokinase/homoserine dehydrogenase 1, chloroplast
 DE precursor (AK-HD 1) (AK-HSDH 1) [includes: Aspartokinase (EC 2.7.2.4);
 DE Homoserine dehydrogenase (EC 1.1.1.3)].
 GN AKHSDH1.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACMAD clade; Panicoideae; Andropogonaceae; Zea.
 OK NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Seedling leaf;
 RX MEDLINE=95148730; PubMed=7846152;
 RA Mehbauer G.J., Somers D.A., Matthews B.F., Gengenbach B.G.;
 RT "Molecular genetics of the maize (Zea mays L.) aspartate kinase-
 RT homoserine dehydrogenase gene family.";
 RL Plant Physiol. 106:1303-1312(1994).
 CC -1- CATALYTIC ACTIVITY: L-homoserine + NAD(P) (+) = L-aspartate 4-
 CC semialdehyde + NAD(P)H.
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
 CC aspartate.
 CC -1- PATHWAY: Catalyzes 2 nonconsecutive reactions in the common
 CC biosynthetic pathway leading from Asp to diaminopimelate and Lys,
 CC to Met, and to Thr and Ile.
 CC -1- SUBUNIT: Homo- or heterodimer (Potential).
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: In the N-terminal section; belongs to the
 CC aspartokinase family.
 CC -1- SIMILARITY: In the C-terminal section; belongs to the homoserine
 CC dehydrogenase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL, L33912; AAA74360.1; --
 DR PIR; T02953; T02953.

DR MaizeDB; 66609; --
 DR InterPro; IPR001048; Aa_kinase.
 DR InterPro; IPR002912; ACT.
 DR InterPro; IPR001341; Aspartate_kinase.
 DR InterPro; IPR001342; Homoserine_dh.
 DR InterPro; IPR005106; NAD_binding_3.
 DR Pfam; PF00696; aak_kinase; 1.
 DR Pfam; PF01842; ACT; 2.
 DR Pfam; PF00742; Homoserine_dh; 1.
 DR Pfam; PF03447; NAD_binding_3; 1.
 DR TIGRfams; TIGR00657; asp_kinase; 1.
 DR PROSITE; PS00324; ASPARTOKINASE; 1.
 DR PROSITE; PS01042; HOMOSER DHEGENASE; 1.
 DR Transferrase; Kinase; Oxidoreductase; Methionine biosynthesis; NADP;
 KW Multifunctional enzyme; Chloroplast; Transient peptide;
 KW Multigene family.
 KM TRANSIT
 FT CHAIN 1 92 CHLOROPLAST (POTENTIAL).
 FT TRANSIT 93 92 BIFUNCTIONAL ASPARTOKINASE/HOMOSERINE
 FT TRANSIT 93 92 DEHYDROGENASE 1.
 FT DOMAIN 93 341 ASPARTOKINASE.
 FT DOMAIN 342 566 INTERFACE.
 FT DOMAIN 567 920 HOMOSERINE DEHYDROGENASE.
 FT NP BIND 568 573 NADP (POTENTIAL).
 FT SEQUENCE 920 AA; 100335 MW; 08DCF444BE645529 CRC64;
 Query Match 16.1%; Score 452.5; DB 1; Length 920;
 Best Local Similarity 26.9%; Pred. No. 8,7e-24;
 Matches 154; Conservative 115; Mismatches 239; Indels 65; Gaps 20;
 QY 8 AAAPRLVPSIPPPASSGHRGL-ACFGTRGPGARGLSVNVADSTRRAKQADGGDV- 65
 DB 13 AAFSTRPRLPHPPAAGDSTFGRCWEKTCQSSFESSL-----RISRLPRTYH-GDILK 67
 QY 66 -LGAVLGLGMEGL-----GDQLSVMKFGSGSVSSAARVAEYAGLITFPERPV 116
 DB 68 NLARTPAAGVEQAEALADLPKDDMS-VHKFGTCMGSEKIHNVADIVLADPBRKL 126
 QY 117 VLSMGKTNNLLIAGKAVGCVITVSELENNMKSLHITVDEL-----GLPKI 169
 DB 127 VVVSAMSKYTTDMVNLVKAQSRDSTYAVLDE---VFDKMTYADKLLAGEDLAPLSQ 183
 QY 170 CNTSLYELIOLKIGIAMKELTPRTDVLVSGECWSTRIIPSAVINKIRYKAOYADF 229
 DB 184 LHADISNLKAMIRAIYIGHATESPDSFVGHGSLMAQULSVAIQSGTPCSMDTREV 243
 QY 230 GFTTDERGNADILEATYPAVAKLHGDWITODPAIPV-TGFLGKWKSGAVTTLGSGS 288
 DB 244 LVVNP---SGANQVDPDYLESEKLEKFSRCPAETIIATGFIH-STPENIPTTLKRDG 299
 QY 289 DLTATTTGKALGLREIQWVDVGLTODPNITPHAKTIPVLTPEEATELAFGAQVLP 348
 DB 300 DPSAIIIGSLVAKRVITMTDVGFSADPRKSEAVITLSTISQEMWMSYGCANVLP 359
 QY 349 QSMRPARBGDIPIRVKNSYNPKAPGTLITRQ----RDMXGLVLTSTIVLKNVTMDI 403
 DB 360 RTIIPVKXNIPDIVRNIFNTSAPGTWICQPANENGLECYAFAPI--DKIALVNV 416
 QY 404 VETRMIGQGYFLARVSGICIEDLCISYDVC--ATSEVSVSLDPSKISRRELIQOASE 461
 DB 417 EGTGAAGVGTANALFGA--VKDVGANTIMISQASSEHSVCFAV-PEK---EVALVSA 469
 QY 462 LDHVEE--LEKIAIVRLLOQRAIISLIQVNEQSSLILEKTGRV-----LRKSGNV 511
 DB 470 LHAAREALAGRLSKVEIVHNCSILATVG-----LRMASTPGVSAITLPDALAKANIN 523
 QY 512 QMTSQASKVNSLIYHSDAKALVBAHQAF 544
 DB 524 RAIAQGCSEYNTIVLKQDCVRAIIRAHSRFF 556
 RESULT 10
 AKH1_BUCBP
 ID AKH1_BUCBP STANDARD; PRT; 816 AA.

AC Q99AR4; 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Bifunctional aspartokinase/homoserine dehydrogenase (AK-HD)
 DE [includes: Aspartokinase (EC 2.7.2.4); Homoserine dehydrogenase (EC 1.1.1.3)]
 GN THKA OR BHP183;
 OS Buchnera aphidicola (subsp. *Baizongia pistaciae*).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Buchnera.
 OX NCBI_TaxID=135942;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426901; PubMed=12522265;
 RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,
 RA Bactolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,
 RA Tames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.;
 RA "Reductive genome evolution in *Buchnera aphidicola*."
 RT Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).
 CC -1- CATALYTIC ACTIVITY: L-homoserine + NAD(P)(+) = L-aspartate + 4-semialdehyde + NAD(P)H.
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-aspartate.
 CC -1- PATHWAY: Catalyzes 2 nonconsecutive reactions in the common biosynthetic pathway leading from Asp to diaminopimelate and Lys, to Met, and to Thr and Ile.
 CC -1- SUBUNIT: Homotetramer (By similarity).
 CC -1- SIMILARITY: In the N-terminal section; belongs to the aspartokinase family.
 CC -1- SIMILARITY: In the C-terminal section; belongs to the homoserine dehydrogenase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF014016; AAC06915.1; -
 DR InterPro; IPR001048; Aa_kinase.
 DR InterPro; IPR001341; Aspartate_kinase.
 DR InterPro; IPR001342; Homoserine_dh.
 DR InterPro; IPR005106; NAD_binding_3.
 DR Pfam; PF00696; aakinese; 1.
 DR Pfam; PF00742; Homoserine_dh; 1.
 DR Pfam; PF03447; NAD_binding_3; 1.
 DR PROSITE; PS00324; ASPARTOKINASE; 1.
 DR PROSITE; PS01042; HOMOSER_DHGENSE; 1.
 DR TRANSFERASE; Kinase; Oxidoreductase; Threonine biosynthesis; NADP;
 KM Multifunctional enzyme; Complete proteome.
 FT DOMAIN 1 250 ASPARTOKINASE (BY SIMILARITY).
 FT DOMAIN 251 471 INTERPACE (BY SIMILARITY).
 FT DOMAIN 472 816 HOMOSERINE DEHYDROGENASE (BY SIMILARITY).
 FT NP_BIND 472 479 NADP (POTENTIAL).
 FT SEQUENCE 816 AA; 91918 MW; F8C3ADC0114576B5 CRC64;
 SQ
 Query Match 15.6%; Score 440; DB 1; Length 816;
 Best Local Similarity 26.8%; Pred. No. 5.5e-23;
 Matches 141; Conservative 96; Mismatches 182; Indels 108; Gaps 16;
 QY 86 VMKRGSSVSSAKRAAEVAGLI-LTFPERPVYLTSMGKTNNLLAGSKA-GGGVH 143
 DB 3 ILKRGSTLSNSELFEHVAITIEENNINNEQTAIVLSPGNTNLLSLAINQTNKKXILP 62
 QY 144 VSEIEENNVKSLH-IKTVDELGLFXICN--TSLVELRLKGIAMKELPPRSDYI 198
 DB 63 IVQKIEKPKLKNINIVGEQKLYEKIKNNIEKMLKLNKILGIVLWLCQCPDKIRAKI 122
 QY 199 VSFEGCMSTRISAVLNKIRKARQYDAFDIGFTT-----DEFGNADILERT 246

DB 123 ISSGEVLSISINSL-----ISGVNTTIIDPVKLLTKEDTYLNATV----- 166
 QY 247 YPAVAK-RLHGMIODDPAIPVVTGLGKMGSGAVTTIGRGSDLTATTICKALGLEIQ 305
 DB 167 NIKISFRILSKMIPKPHIILMPGFTA-GNKGELVTLGRNGSDYSATILSVCLNSTWCE 225
 QY 306 VWRDVGVLTCDPNIYPAKTVPLYTFEATELAYFGAQQVAPQSMRPARPGDIPVRVKN 365
 DB 226 IWDVANGVITCPKLVSDAKLTLSTSYNEALISTLGLKILHPNTIYIQKFKIPCTKN 285
 QY 366 SYNPKAPGTLITQRDMXGLVLTSLVKSNTWMLDVSTRMLQGYFLARVSGICIE 425
 DB 286 THNPSSIGTKISCNHVNKNL-----ITGVYLE 314
 QY 426 DL-CIVDCVANEVSIVS-----SLDPSKIMRELIOAS----- 460
 DB 315 NVNMFSTICLSKNITITIPKIFSCMSLSKIIITITQTSNTISPCILKMTNTALHV 374
 QY 461 -----ELDHVELEKIAIVLLQORAIISLVNVEQSSIIIEKTRVIRKSGVNYQM 513
 DB 375 LHKALVDELGRKL-LKPIKVERKULLISVIS--SDIINNTKTERKVSILKHWVINTLA 430
 QY 514 ISGASAKVMSLIVHSDPAKALVELHQAFFEDVDVLSQVEANLLVG 560
 DB 431 ISKGASKNISIVVKHDDILGVRLHKIEFN-----KQCTAIFLIG 473
 RESULT 11
 ID AK_THETH STANDARD; PRT; 405 AA.
 AC P97151; F77991;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Aspartokinase (EC 2.7.2.4) (Aspartate kinase) [contains: Aspartokinase alpha subunit (ASK-alpha); Aspartokinase beta subunit (ASK-beta)].
 GN ASK OR ASKAB.
 OS Thermus thermophilus.
 OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
 OC Thermus.
 OX NCBI_TaxID=274;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=HB27;
 RX MEDLINE=99173891; PubMed=10074061;
 RX Kobashi N., Nishiyama M., Tanokura M.;
 RA "Aspartate kinase-independent lysine synthesis in an extremely thermophilic bacterium, *Thermus thermophilus*: lysine is synthesized via alpha-aminoadipic acid not via diaminopimelic acid."
 RT J. Bacteriol. 181:1713-1718(1999).
 RT [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 33923 / AT-62;
 RX MEDLINE=95291465; PubMed=773416;
 RA Nishiyama M., Kukimoto M., Beppu T., Horiuchi S.;
 RA "An operon encoding aspartokinase and purine phosphoribosyltransferase in *Thermus flavus*."
 RT Microbiology 141:1211-1219(1995).
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-aspartate.
 CC -1- PATHWAY: First step in the common biosynthetic pathway leading from Asp to diaminopimelate, to Met, to Ile and to Thr.
 CC -1- SUBUNIT: Tetramer consisting of 2 isoforms Alpha (catalytic) and 2 isoforms Beta (function not known) (By similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event:Alternative initiation:
 CC Comment=2 isoforms, Alpha/Aspartokinase alpha subunit (shown here) and Beta/Aspartokinase beta subunit, may be produced by alternative initiation;
 CC -1- SIMILARITY: Belongs to the aspartokinase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB013311; BAA25848.1; -
 DR EMBL; AB013311; BAA25849.1; -
 DR EMBL; D37928; BAA07146.1; -
 DR EMBL; D37928; BAA07147.1; -
 DR InterPro; IPR001048; Aa_kinase.
 DR InterPro; IPR002912; ACT.
 DR InterPro; IPR005260; Asp_kin_monofn.
 DR InterPro; IPR001341; Aspartate_kinase.
 DR Pfam; PF00696; aak_kinase; 1.
 DR Pfam; PF01842; ACT; 2.
 DR TIGRPFAMs; TIGR00656; asp_kin_monofn; 1.
 DR TIGRPFAMs; TIGR00657; asp_kinase; 1.
 DR PROSITE; PS00324; ASPARTOKINASE; 1.
 KW Transferase; Kinase; Diaminopimelate biosynthesis;
 KW Lysine biosynthesis; Alternative initiation.
 FT CHAIN 1 405 ASPARTOKINASE, ISOFORM ALPHA.
 FT CHAIN 245 405 ASPARTOKINASE, ISOFORM BETA.
 FT INIT MET 245 245 FOR ISOFORM BETA.
 FT VARIANT 126 126 E -> D (IN STRAIN AT-62).
 SQ SEQUENCE 405 AA; 43333 MW; 06F7B79E9C8960EA CRC64;
 Query Match 14.8%; Score 418; DB 1; Length 405;
 Best Local Similarity 27.9%; Pred. No. 7,3e-22;
 Matches 135; Conservative 88; Mismatches 149; Indels 112; Gaps 18;
 QY 85 VVMKFGSSVSAAAMAEVAGLILTPPE-RRPVVLASMGKTNNLLAGEKAVGCVI 142
 DB 4 VVQKYGSTVGDLEHKKVQRIAHYREKGRHLAVVVSAMGHT----- 47
 QY 143 HVSEIEEMNWKSLHKTVDLGLPXCNTSLYELQQLKGIAMKELTPRTS----DYL 198
 DB 48 -----DEL-----TALKRVNPRPPEDLL 69
 QY 199 VSPGECWSTRIFSAVLNKKIPKARQYDADPFGITTFDEFGNADILEATPAVAKRIHDM 258
 DB 70 TTTCGEYSVALLMQMLWAGIPAKGVVQDGIITDGRGARDLILEVNPARIREL---- 125
 QY 259 IQDDALPVVTGPIGKWKSGAVTTLRGSSDLTATTIGALGLREIQWVKDVGVLCDP 318
 DB 126 -EQGFVAVIAGFMGT-TPEGEITTLGRGSDTTVAIAALAKACEIYTTDEGYTTDP 183
 QY 319 NIYPKATVPYLTFFEEATELAFGAOVLPQSMRPARBGDIPRVKNSYNPKAPGTLITR 378
 DB 184 HILPEARKLSVIGDQMLEMAAGAVLHPRVYAKRGVLAHRSSTYN-PETLVKE 242
 QY 379 -QSDMDXGLVVLTSIVLKSNTMLDIVSRMIG--QYGFLLARVSGICYEDLCISVDCV 434
 DB 243 VVAMEMK--AVTGVALD-----LDHAQIGLIGIPDPPIAAKV-----FQ 280
 QY 435 ATSEVSVSVSL-----DPK-----IWSRELIOQASE-LDHVEELEKIAIVRLIQ 480
 DB 281 ALMERGLAVMTIQGVGADPSQQAFAFYKCDPQAEALALEPLAIEGSAIIR--PD 338
 QY 481 RAIIISLIG-NVEOSSILEKTVLRKSGVNVQMTISQASKVNMSLIVHSDAKALVEAL 539
 DB 339 IAKYSIVGVGLASTPEVPAPMFQAVASTGANIEMIA--TSEVRISVIIIPAEVAAALRAV 366
 QY 540 HQAF 543
 DB 397 HQAF 400

DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Aspartokinese (EC 2.7.2.4) (Aspartate kinase) [Contains: Aspartokinese
 DE alpha subunit; Aspartokinese beta subunit].
 GN LYSC OR ASK OR CBO220.
 OS Corynebacterium efficiens.
 CC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 CC Corynebacterinae; Corynebacteriaceae; Corynebacterium.
 CX NCBI_TaxID=152794;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
 RA Itaya H., Kimura E., Kawahara Y., Sugimoto S.;
 RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
 RA Gojohori T.;
 RT "Comparative complete genome sequence analysis of the amino acid
 RT replements responsible for the thermostability of Corynebacterium
 RT efficiens.";
 RL Genome Res. 13:1572-1579(2003).
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
 CC aspartate.
 CC -1- ENZYME REGULATION: Feedback inhibition by lysine and threonine.
 CC -1- PATHWAY: First step in the common biosynthetic pathway leading
 CC from Asp to the cell wall precursor meso-diaminopimelate, to Lys,
 CC to Met, to Ile and to Thr.
 CC -1- SUBUNIT: Tetramer consisting of 2 isoforms Alpha (catalytic) and 2
 CC isoforms Beta (function not known).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative initiation;
 CC Comment=2 Isoforms, Alpha/Aspartokinese alpha subunit (shown
 CC here) and Beta/Aspartokinese beta subunit, are produced by
 CC alternative initiation;
 CC -1- SIMILARITY: Belongs to the aspartokinese family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB083129; BAA88820.1; -
 DR EMBL; AP005214; BAC17030.1; ALU_INIT.
 DR InterPro; IPR001048; Aa_kinase.
 DR InterPro; IPR002912; ACT.
 DR InterPro; IPR005260; Asp_kin_monofn.
 DR InterPro; IPR001341; Aspartate_kinase.
 DR Pfam; PF00696; aak_kinase; 1.
 DR Pfam; PF01842; ACT; 2.
 DR TIGRPFAMs; TIGR00656; asp_kin_monofn; 1.
 DR TIGRPFAMs; TIGR00657; asp_kinase; 1.
 DR PROSITE; PS00324; ASPARTOKINASE; 1.
 KW Transferase; Kinase; Diaminopimelate biosynthesis;
 KW Lysine biosynthesis; Alternative initiation; Complete proteome.
 FT CHAIN 1 421 ASPARTOKINASE, ISOFORM ALPHA.
 FT CHAIN 250 421 ASPARTOKINASE, ISOFORM BETA.
 FT INIT MET 250 250 FOR ISOFORM BETA.
 SQ SEQUENCE 421 AA; 44793 MW; F0D63830E93CB5F CRC64;
 Query Match 14.8%; Score 416.5; DB 1; Length 421;
 Best Local Similarity 28.1%; Pred. No. 5.9e-22;
 Matches 134; Conservative 88; Mismatches 162; Indels 93; Gaps 14;
 QY 85 VVMKFGSSVSAAAMAEVAGLILTPPE-RRPVVLASMGKTNNLLAGEKAVGCVI 142


```

Cc 4 VVKRYGSSLESERIRNVAERIVATKACNDVVVCSMGTTBELL----- 51
Cc 143 HVEIEEMWVKSIIKIVDELGFXICNTSLYEELQLLKGIAMKELTPRTSDYVSG 202
Cc 52 -----DLAAVNPVPAPREM-----DMLITAG 73
Cc 203 ECKSTRIFPAVINKIRKARQYAPDIFGITTEDEGNMILATPAVAKLHGMWDG 262
Cc 74 ERTSNALVAVALSGLGASQSTFGSAGVLTTERGNARIVVTGRTREAL-----DEG 128
Cc 263 AIPVYTGKCKGKSGAVTTLGSGSDLTATTIGKALGRIEQWADVGVLTCDPNITP 322
Cc 129 KICIVAGFGQVKNKEDVTTLTGSGSDTLTAVLAALADGVCIEISVDGVYADPRIVP 188
Cc 323 HACTVYLTTEETELAVGAGVLTQHSRPARREGIIPRVKNSYNPKAPGLTIR-RQRD 381
Cc 189 NAQKLERLSEEWLELAAGSKTLVRSVEYAAFAVPRVRSYS-NDPGTLIGSWED 247
Cc 382 MDXGLVLTSTVL---KSNVTMLDIVSTMLGQYGLFARVSGICIEDLCISVDCV--- 434
Cc 248 IPWEAVLTGAVATDKSEAKVTVGIDP-----KPGAAKV--FRALADAEIWDVLTQNV 300
Cc 435 -----ATSEVSVV--SLDPSKWSRELIQASSELNHYVELEKAIYRLLOQRAIILI 487
Cc 301 SSVEDGTTDITFTCPRSDGPRAMELKRYQOQGDWNTVLYD-----DQVGKSVLV 350
Cc 488 G-NVEQSLILEKTRVLRKSGVNVQMGASKVMSLIVHSDPAKLVLEALHQA 543
Cc 351 GAGKSHPVTAFFMALRPNVNVVLLIS--TSEIRISVLIREDLDSAKALHKEF 405

RESULT 13
AKH_BUCAP STANDARD; PRT; 814 AA.
ID AKH_BUCAP STANDARD; PRT; 814 AA.
AC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bifunctional aspartokinase/homoserine dehydrogenase (AK-HD)
DE [includes: Aspartokinase (EC 2.7.2.4); Homoserine dehydrogenase (EC 1.1.1.3)].
CN THRA OR BUSC188.
OS Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=98794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22084549; PubMed=12089438;
RA Tamás I., Klasson L., Canhaeck B., Naeslund A.K., Eriksson A.-S.,
RA Wernegreen U.J., Sandström J.P., Moran N.A., Andersson S.G.E.,
RA "50 million years of genomic stasis in endosymbiotic bacteria.";
RL Science 296:2376-2379 (2002).
CC -1 CATALYTIC ACTIVITY: L-homoserine + NAD(P)(+) = L-aspartate 4-
CC semialdehyde + NAD(P)H.
CC -1 CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
CC aspartate.
CC -1 PATHWAY: Catalyzes 2 nonconsecutive reactions in the common
CC biosynthetic pathway leading from Asp to diaminopimelate and Lys,
CC to Met, and to Thr and Ile.
CC -1 SUBUNIT: Homotetramer (By similarity).
CC -1 SIMILARITY: In the N-terminal section; belongs to the
CC aspartokinase family.
CC -1 SIMILARITY: In the C-terminal section; belongs to the homoserine
CC dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/

```

```

Cc or send an email to license@isb-sib.ch).
Cc EMBL AB014094; AM67753.1; -.
Cc InterPro: IPR001048; AA_kinase.
Cc InterPro: IPR002912; ACT.
Cc InterPro: IPR001341; Aspartate_kinase.
Cc InterPro: IPR001342; Homoserine_dh.
Cc InterPro: IPR005106; NAD_binding_3.
Cc Pfam: PF00696; aak_kinase; 1.
Cc Pfam: PF01842; ACT; 2.
Cc Pfam: PF00742; Homoserine_dh; 1.
Cc Pfam: PF03447; NAD_binding_3; 1.
Cc TIGRfam: TIGR00657; asp_kinases; 1.
Cc PROSITE: PS00324; ASPARTOKINASE; 1.
Cc PROSITE: PS01042; HOMOSER DHGENASE; 1.
Cc K0 Transferrase; Kinase; Oxidoreductase; Threonine biosynthesis; NADP;
Cc Multi-functional enzyme; Complete proteome.
Cc FT DOMAIN 1 248 ASPARTOKINASE (BY SIMILARITY).
Cc FT DOMAIN 249 469 INTERPACE (BY SIMILARITY).
Cc FT DOMAIN 470 469 HOMOSERINE DEHYDROGENASE (BY
Cc SIMILARITY).
Cc FT NP_BIND 470 477 NADP (POTENTIAL).
Cc SQ SEQUENCE 814 AA; 91378 MW; 162CDD52590D1A00 CRC64;

Query Match 14.8%; Score 416.5; DB 1; Length 814;
Best Local Similarity 25.7%; Pred. No. 2,4e-21;
Matches 127; Conservative 111; Mismatches 197; Indels 59; Gaps 14;

Cc 86 VMKFGSSVSAAAEVAGLI-LTPPEPRPVVLTAMGKTNNLLAGEKAVGCVIHW 144
Cc 3 LKFGSTLANAKKFLCVADIEKKKKEQIAVVASPAKITNTVLTATIEKKIDDEV--- 59
Cc 145 SEIEEMWVKSIIKIVDELG-----LXICHTSYLE-----QLKGIAMKELTPRT 194
Cc 60 --LKKINLAKNFIIEIDQIKRIQPLPYENTKSTIELEPKKIINGILLIQCPEGI 117
Cc 195 SDYVSEFGMSRTIFSAVINKIRKARQYADFDIGITTT-DEFGN-----ADILEATYP 248
Cc 118 KPITIRGELISVDIMKIL-----QSNHEVTILNPNTNLISGNYLDSITDIKES--- 169
Cc 249 AVAKRLHGMWIDPAIPVVTGVLGKGMWGAATTLGSGSDLTATTIGKALGLEIOWK 308
Cc 170 --KKRIKKINIDQKNIIMAGFIA-GNKEGEJLVVGRNGSYSAIIASCNAKCEITWT 226
Cc 309 DVDGVLTCDPNITPYPAKTVPYLTPEARELAYAGVAVHPOSMPAREGDI-PVAVKNSYN 368
Cc 227 DVDGVLTADPRIVSNLYLDIYSQDAHELSPFAKVLHPTIETISQFQIPCVIKNNN 286
Cc 369 PKAPGLITRQDMXGLVLTSTVLSKSNVTMLDIVSTRMLGQYGLFARVSGICIEDLC 428
Cc 287 TESKGTWIGKNNPNDN--SLKGYVYLDIIMFNISGCLKDSGNTIARIFTILSREBMK 344
Cc 429 ISVDCAVATSEVSVSLDPSKWSRELIQASSELNHYVELEK-----IAIVLL 478
Cc 345 ILLIIOSSSENQVN-----FCTFEKDIYILLILKETTETIKESGLNDPRIV 392
Cc 479 QORAILSLIG-NVEQSLILEKTRVLRKSGVNVQMGASKVMSLIVHSDPAKLVLE 537
Cc 393 KNLTILISVGSNISSEKNTASKIFSSLGSKINVLAIAGSSKHSISIVIKKENLLQGIQ 452
Cc 538 ALHQA-PPEDVUS 550
Cc 453 NINHTLFFKXTIIN 466

RESULT 14
AK_CORGL STANDARD; PRT; 421 AA.
ID AK_CORGL STANDARD; PRT; 421 AA.
AC P26512; Q59286;
DT 01-AUG-1992 (Rel. 23, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Aspartokinase (EC 2.7.2.4) (Aspartate kinase) [contains: Aspartokinase

```

DE alpha subunit; Aspartokinase beta subunit].
GN LYSC OR CGJ0251.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX MEDLINE=92065816; PubMed=1956296;
RA Kainowski U., Cremer U., Bachmann B., Eggeling L., Sahn H.,
RA Puhler A.,
RT "Genetic and biochemical analysis of the aspartokinase from
RT Corynebacterium glutamicum";
RL Mol. Microbiol. 5:1197-1204(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.,
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032";
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-51 FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX MEDLINE=94161495; PubMed=8117072;
RA Patek M., Kumbach K., Eggeling L., Sahn H.,
RT "Leucine synthesis in Corynebacterium glutamicum: enzyme activities,
RT structure of leuA, and effect of leuA inactivation on lysine
RT synthesis";
RL Appl. Environ. Microbiol. 60:133-140(1994).
RN [4]
RP SEQUENCE OF 158-421 FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX MEDLINE=91094767; PubMed=1980002;
RA Kainowski U., Bachmann B., Thierbach G., Puhler A.,
RT "Aspartokinase genes lysC alpha and lysC beta overlap and are
RT adjacent to the aspartate beta-semialdehyde dehydrogenase gene and in
RT Corynebacterium glutamicum";
RL Mol. Gen. Genet. 224:317-324(1990).
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
CC aspartate.
CC -1- ENZYME REGULATION: Feedback inhibition by lysine and threonine.
CC -1- PATHWAY: First step in the common biosynthetic pathway leading
CC from Asp to the cell wall precursor meso-diaminopimelate, to Lys,
CC to Met, to Ile and to Thr.
CC -1- SUBUNIT: Tetramer consisting of two isoforms Alpha (catalytic) and
CC two isoforms Beta (function not known).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative initiation;
CC Comment=2 isoforms, Alpha/Aspartokinase alpha subunit (shown
CC here) and Beta/Aspartokinase beta subunit, are produced by
CC alternative initiation;
CC -1- SIMILARITY: Belongs to the aspartokinase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: X5726; CAA40502.1; -
CC EMBL: X5726; CAA40503.1; -
CC EMBL: AP005274; BAB97644.1; -
CC EMBL: X70959; CAA50296.1; ALT_SEQ.
CC PIR: 140723; 140723.
CC PIR: S15276; S15276.
CC InterPro: IPR001048; Aa_kinase.
CC InterPro: IPR002912; ACT.
CC InterPro: IPR005260; Asp_kin_monofn.
CC InterPro: IPR001341; Aspartate_kinase.
CC Pfam: PF00696; aakinae; 1.

DR Pfam: PF01842; ACT; 2.
DR TIGRFAMs: TIGR00656; asp_kin_monofn. 1.
DR TIGRFAMs: TIGR00657; asp_kinae; 1.
DR PROSITE: PS00324; ASPARTOKINASE; 1.
KW Transferase; Kinase; Diaminopimelate biosynthesis; Complete proteome.
KW Lysine biosynthesis; Alternative initiation; Complete proteome.
FT CHAIN 1 421
FT INIT MET 250 421 ASPARTOKINASE, ISOFORM BETA.
FT CHAIN 250 421 FOR ISOFORM BETA.
FT MUTAGEN 301 301 S->Y: FEEDBACK-RESISTANT AND ENHANCED
FT CONFLICT 40 40 EXPRESSION OF THE ASD GENE.
SQ SEQUENCE 421 AA; 44754 MW; E36BAD0061D50827 CRC64;
C -> V (IN REF. 1 AND 3).
Query Match 14.8%; Score 415.5; DB 1; Length 421;
Best Local Similarity 29.0%; Pred. No. 1.2e-21;
Matches 137; Conservative 84; Mismatches 167; Indels 85; Gaps 14;
QY 85 VMKRGSSVSAAEMAEVAGLITPEE--ERPVTLSAMGKTTNNLLAGEKAVGCVI 142
DB 4 VVQKGSLSAERINVAERIVATKAANDVVCASMGDTDBL----- 51
QY 143 HVSEIEEMNVKSLHIKTVDLGLPKICNTSLVELQLKGIAMKELTPRTSDYVSTG 202
DB 52 -----ELAAVVPVPAREM-----DMLLTAG 73
QY 203 EGMSTRISAVINKIRKARQYDAFDIGFITTDFGADILFATYFAVKRLHGDWIODP 262
DB 74 ERISALVAKMAIESLGAEDSFTGSQAGVLFTEHGRARIVDTVPGVREAL-----DEG 128
QY 263 AIPVVTGFLGKMGSGVATTLGRGSDLTATIGKALGREIWMKDVGVLTCDNITP 322
DB 129 KICIVAGFGVNNKTRDTTLGRGSDTLVALAALNADVCIEYSDVGVYADPRIVP 188
QY 323 HAKTVPYLTPEATELAFGAQVILHPSGMAREGDIIPVVKSNYPKAPGLIT-RQD 381
DB 189 NAKTKETKSFEMELAAVGSKILVLSVEYAFNPLVRSYS-NDPGLILASGMD 247
QY 382 MDXGLVLTSLVL--KSNVTMLDIVSTRMLQGYGLARYSGICYIEDLCISVQVATSE 438
DB 248 IPEEAVLTGVAITKSEKVTVLGISD-----KGEAAKV--PALADAEINIDWLVQNV 300
QY 439 VSV-SVSDPSKINSRELIOASLDHVELEKIAL-----VRLIQGAILSLG-NV 490
DB 301 SSEVDGTDITFTCPERSDGRAMEI-----LKKLQVGWNTVLVDDQGVKSLVAGM 354
QY 491 EGSLLKTRGVLRKSGVNYQMSQASKVNLSLIYHSDAKALVEALHQA 543
DB 355 KHPGVTAEPHEALRDVNVNIELIS--TSEIRISVLREDDLDAAARALHECF 405
RESULT 15
AK MYCSM STANDARD; PRT; 421 AA.
ID AK MYCSM
AC P41403;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Aspartokinase (EC 2.7.2.4) (Aspartate kinase) [Contains: Aspartokinase
DE alpha subunit (ASK-alpha); Aspartokinase beta subunit (ASK-beta)].
GN ASK.
OS Mycobacterium smegmatis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1772;
OX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 607 / mc(2)6 / NRRL B-692;
RX MEDLINE=94254720; PubMed=7910936;
RA Cirillo J.D., Weisbrod T.R., Pascopella L., Bloom B.R.,
RA Jacobs W.R. Jr.,
RT "Isolation and characterization of the aspartokinase and aspartate
RT semialdehyde dehydrogenase operon from mycobacteria";

```

RL Mol. Microbiol. 11:629-639(1994).
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
CC aspartate.
CC -1- ENZYME REGULATION: Feedback inhibition by lysine and threonine (By
CC similarity).
CC -1- PATHWAY: First step in the common biosynthetic pathway leading
CC from Asp to the cell wall precursor meso-diaminopimelate, to Lys,
CC to Met, to Ile and to Thr.
CC -1- SUBUNIT: Tetramer consisting of 2 isoforms Alpha (catalytic) and 2
CC isoforms Beta (function not known) (By similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative initiation;
CC Comment=2 isoforms: Alpha/Aspartokinase alpha subunit (shown
CC here) and Beta/Aspartokinase beta subunit, may be produced by
CC alternative initiation;
CC -1- SIMILARITY: Belongs to the aspartokinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z17372; CAA78984.1; -.
DR EMBL: Z17372; CAA78985.1; ALT_INIT.
DR PIR: S42422; S42422.
DR InterPro: IPR001046; Aa_kinase.
DR InterPro: IPR002912; ACT.
DR InterPro: IPR005260; Asp_kin_monofn.
DR InterPro: IPR001341; Aspartate_kinase.
DR Pfam: PF00693; aak:kinase; 1.
DR Pfam: PF01842; ACT; 2.
DR TIGRfams: TIGR00636; asp_kin_monofn; 1.
DR TIGRfams: TIGR00657; asp_kinases; 1.
DR PROSITE: PS00324; ASPARTOKINASE; 1.
DR Transferase; Kinase; Diaminopimelate biosynthesis;
DR Lysine biosynthesis; Alternative initiation.
FT CHAIN 1 421 ASPARTOKINASE, ISOFORM ALPHA.
FT INT MET 250 421 ASPARTOKINASE, ISOFORM BETA.
FT INT MET 250 250 FOR ISOFORM BETA.
SQ SEQUENCE 421 AA; 44458 MW; 926B9FC62520CD CRC64;

Query March 14.8%; Score 415.5; DB 1; Length 421;
Best Local Similarity 27.5%; Pred. No. 1.2e-21;
Matches 134; Conservative 87; Mismatches 152; Indels 115; Gaps 16;

QY 85 VMKFGSSVSAAKMAEVAAGLITPE--ERPVTLSMGKTTNNLLAGEKAVGCVI 142
DB 4 VVQKYGSSVADERIRVAERIVETKAGNDVVVVASAGDTIDLL----- 51
QY 143 HVSEIEMNWKSLHITVDLGLPXICNTSLYEQLKGLAMKELTPRTSDYLVSPG 202
DB 52 -----DLARQVSPAP--PPREMDMLTFAG 73
QY 203 ECMSTRLFSAYLTKIRKARQYDAFDIGTTDFGNAIDLENTYPAVAKRLHGMIDP 262
DB 74 ERIENALVMAIEISLGAQARSFTGSCAGVITTTGHNAKIIDVT---PGLRL-DALDEG 128
QY 263 AIPVVTGLGKWSGAVTTLGRGSDLTATTIGKALGLEIOWMKVDVGLTCDENIYP 322
DB 129 QIVVAGFGVSQSKVTLTLGRGSDTLTAVAALADAVCEIYTDVGIPTADPRIVP 188
QY 323 HAKTVPYLTFFEEATELAYFGAQVLAHQSMRPARBGDIPVRKNSYNPKAPGTLITRQ-RD 381
DB 189 NARHLDTVSFEEMLEMAACGAKVLMRCVEYARRYNPDIHRSYSYDK--FGTIIVKGSIED 247
QY 382 MDXGLVLTSTV--LTSNTMDIVSTRMLGQVFLARVSGICYIEDLCISVDCAVATSE 438
DB 248 IPMEDAILGVAAHRSSEAKTVVGLPDVP-----GYAAKYFR-----AVAE 288
QY 439 VSVSVSLDPSKIMRELIQASELD-----HVVEEL---EKIAIVRLI 478

```

```

DB 289 ADVNIDM-----VLQNSKIEDKTDITFCARDNBPRAVERKLSALKSEIGFQVL 339
QY 479 QQRAL--ISLGNVEQSLITEKT-GRVLRKSSGVNVQMISQASKVMSLIYHDSDAKL 535
DB 340 YDDHIGKYSLLIGAGWRSHPGVTATFCALAEAGINIDLS--TSEIRISVLINQTELDKA 397
QY 536 VEALHQAF 543
DB 398 VGLHBAF 405

```

Search completed: March 23, 2004, 13:16:13
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 13:00:27 ; Search time 71 Seconds
(without alignments)
2468.594 Million cell updates/sec

Title: US-09-890-813-6

Perfect score: 2813
Sequence: 1 MALPVRSAAPRLVPSTIP.....QAFEDVLSQVEANILVWG 560

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1881	66.9	510	10	Q8LPZ8
2	1880	66.8	565	10	Q851Z6
3	1868.5	66.4	581	10	Q8RYL8
4	1835	65.2	564	10	Q9XHC5
5	1789	63.6	569	10	Q9LYU8
6	1784	63.4	569	10	Q23152
7	1747.5	62.1	544	10	Q9FYU4
8	1737.5	61.8	544	10	Q23653
9	1737.5	61.8	544	10	Q9FY44
10	1675.5	59.6	559	10	Q9F702
11	1570	55.8	354	10	Q9MAX0
12	635.5	24.7	470	16	Q8KG73
13	685	24.4	467	17	Q8T230
14	662.5	23.6	479	17	Q8PX05
15	657.5	23.4	472	17	Q8TUD6
16	642.5	22.8	462	17	Q29558

ALIGNMENTS

RESULT 1
ID Q8LPZ8 PRELIMINARY; PRT; 510 AA.
AC Q8LPZ8;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 24, Last annotation update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Monofunctional aspartokinase-like protein (EC 2.7.2.4) (Aspartate kinase).
DE OSUNBA0093F16.26.
OS Oryza sativa (japonica cultivar-group).
CC Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
CC Euphorbiaceae; Oryzae; Oryza.
OX NCBI_Taxid=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC clone:OSUNBA0093F16.";
RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSPHO-L-ASPARTATE.
CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.
CC EMBL, AP004332; BAB92806.1; -.
DR GRIMENE; Q8LPZ8; -.
DR GO; GO:0004072; F:aspartate kinase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006652; P:amino acid biosynthesis; IEA.
DR InterPro: IPR001048; Aa_kinase.
DR InterPro: IPR001341; Aspartate_kinase.
DR Pfam; PF00656; aak_kinase; 1.
DR TIGRFAMs; TIGR00657; asp_kinase; 1.
DR PROSITE; PS00324; ASPARTOKINASE; 1.
KW Kinase; Transferase.
SQ SEQUENCE 510 AA; 54670 MW; 12C1603C1DF0D3EE CRC64;

17 607 21.6 473 16 Q9RUL9 Q9RUL9 deinococcus
18 605.5 21.5 818 16 Q8XAX1 Q8XAX1 chlorobium
19 590 21.0 461 16 Q8ZAS1 Q8ZAS1 yersinia pe
20 579.5 20.6 439 16 Q8A7Z9 Q8A7Z9 bacteroides
21 569 20.2 449 16 Q8X5Y1 Q8X5Y1 escherichia
22 569 20.2 449 16 Q831N8 Q831N8 shigella fl
23 566 20.1 449 16 Q8ZK15 Q8ZK15 salmonella
24 566 20.1 449 16 Q8FB45 Q8FB45 escherichia
25 566 20.1 449 16 Q8Z1U8 Q8Z1U8 salmonella
26 550.5 19.6 450 16 Q9CM97 Q9CM97 pasteurella
27 540.5 19.2 451 16 Q8EAC1 Q8EAC1 shewanella
28 534 19.0 450 16 Q87196 Q87196 vibrio para
29 531 18.9 479 16 Q9KUM8 Q9KUM8 vibrio chol
30 519 18.5 811 16 Q8A541 Q8A541 bacteroides
31 517.5 18.4 451 16 Q7VLP7 Q7VLP7 haemophilus
32 500.5 17.8 430 16 Q8DCU6 Q8DCU6 vibrio vuln
33 490.5 17.4 815 16 Q9CPD4 Q9CPD4 pasteurella
34 485.5 17.3 819 16 Q875D0 Q875D0 vibrio para
35 483.5 17.2 825 16 Q9KPK3 Q9KPK3 vibrio chol
36 477.5 17.0 820 16 Q8XAB4 Q8XAB4 escherichia
37 477.5 17.0 820 16 Q7UDU5 Q7UDU5 shigella fl
38 477.5 17.0 834 16 Q83MT0 Q83MT0 shigella fl
39 477.5 17.0 841 16 Q8FLD9 Q8FLD9 escherichia
40 476.5 16.9 820 2 Q8RMX0 Q8RMX0 escherichia
41 476.5 16.9 835 16 Q8P1A8 Q8P1A8 xanthomonas
42 470.5 16.7 835 16 Q8P901 Q8P901 xanthomonas
43 469.5 16.7 820 16 Q8ZS19 Q8ZS19 salmonella
44 469.5 16.7 820 16 Q8ZSR7 Q8ZSR7 salmonella
45 468.5 16.7 454 16 Q8DR20 Q8DR20 streptococcus

Query Match 66.9%; Score 1881; DB 10; Length 510;
 Best Local Similarity 70.0%; Pred. No. 2,66-136;
 Matches 403; Conservative 29; Mismatches 56; Indels 88; Gaps 9;

QY 1 MAIPRSAAAP--RLVPSIPPASSG-----HYRGLACFGRTRGPRGARG 43
 DB 1 MAIALRLAAAPRLRLVSPAPPAPPAIGAGDGRGGGGAARIGVLRACRRRGGGRL 60
 QY 44 LSVVADSTR-PAQADGGGCVGAPVLGSLGHEGLDQLSYMKRGSSVSAARME 102
 DB 61 VAAAAADRSARCRAGV--GAAAAAETLGGIGVG--GDQLSVMKRGSSVSAARME 117
 QY 103 VAGLLTFPEERPVVLLSAMGKTNNLLAGEKAVGCGVHVSEIEEMNVKSLIKITVD 162
 DB 118 VAGLLTFPEERPVVLLSAMGKTNNLLAGEKAVGCGVHVSEIEEMNVKSLIKITVE 177
 QY 163 ELGLP-XICNTSLVLEBOLLKGIAMMKELTPRTSDYLVSGECSTRIFSAYLNKIRVKA 221
 DB 178 ELALPRSVIHTMLDELLEQLKGIAMMKELTRTIDYLVSGECSTRIFAAYLNKIGVKA 237
 QY 222 RQYDAFDIGFITDERGNADILEATYPAVAKRLHGMWIDPAIPVYTGFLGKWKSGAVT 281
 DB 238 RQYDAFDIGFITDDEGNADILEATYPAVAKRLHGMWIDPAIPVYTGFLGKWKSGAVT 297
 QY 282 TLGRGSDLTATTIGKAGLREIQVWKVDGVLTCDPNIYPAKTVPYLTPEEATLAYF 341
 DB 298 TLGRGSDLTATTIGKAGLREIQVWKVDGVLTCDPNIYPAKTVPYLTPEEATLAYF 357
 QY 342 GAGVLAHPSMRPAREGDPYRVKSYNPKAPGTLTRQDMDGVLVTSIVLSKNTML 401
 DB 358 GAGVLAHPSMRPAREGDPYRVKSYNPKAPGTLTRQDMDGVLVTSIVLSKNTML 415
 QY 402 DIVSTRMLQGYGLARVSGICYIEDLCISVDCVATSEVSVSLDPSKIMRELIQOASE 461
 DB 416 DIVSTRMLQGYGLARVSGICYIEDLCISVDCVATSEVSVSLDPSKIMRELIQOASE 471
 QY 462 LDHVELEKIAIVRLLOORAIISLIGNEOSSILLEKTRVLRKSGVNVQMISQASKV 521
 DB 433 LDHVELEKIAIVRLLOORAIISLIGNEOSSILLEKTRVLRKSGVNVQMISQASKV 521
 QY 522 NMSLIHDSDAKALVEALHQAFFEDVLSQVEAENL 557
 DB 472 NMSLIHDSDAKALVEALHQAFFEDVLSQVEAENL 507

RESULT 2
 Q85126 PRELIMINARY; PRT; 565 AA.
 ID Q85126
 AC Q85126;
 DT 01-JUN-2003 (TEMBLrel. 24, Created)
 DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Putative aspartate kinase.
 GN OSJNB0015N08.5.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriactoidae; Oryzae; Oryza.
 NC NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
 RA Overton II L.L., Taitlin T., Kim M.M., Bera J.U., Jin S.S.,
 RA Fiedler D.W., Tallon L.J., Koo H., Zisemann V., Heide J., Blunt S.,
 RA Vanaken S.S., Riedmuller S.B., Uteback T.T., Feldlyum T.V.,
 RA Yang Q.Q., Haas B.J., Sun B.B., Peterson J.U., Quackenbush J.,
 RA White O., Salzberg S.L., Frazer C.M.;
 RT "Oryza sativa chromosome 3 BAC OSJNB0015N08 genomic sequence";
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBD databases.
 DR GO: GO:0016597; F:amino acid binding; IEA.
 DR GO: GO:0004072; F:aspartate kinase activity; IEA.

DR GO: GO:0016301; F:kinase activity; IEA.
 DR GO: GO:0008652; P:amino acid biosynthesis; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR001048; aa_kinase.
 DR InterPro: IPR002912; ACT.
 DR InterPro: IPR001341; Aspartate_kinase.
 DR Pfam: PF00696; aak_kinase; 1.
 DR Pfam: PF01842; ACT; 2.
 DR TIGRFAMs: TIGR00657; aep_kinase; 1.
 KW kinase.
 SQ SEQUENCE 565 AA; 60865 MW; 83103648C6FD7867 CRC64;

Query Match 66.8%; Score 1880; DB 10; Length 565;
 Best Local Similarity 70.7%; Pred. No. 3,66-136;
 Matches 396; Conservative 49; Mismatches 101; Indels 14; Gaps 8;

QY 1 MAIPRSAAAP--RLVPSIPPASSGAVRGLAC-FGRTRGPRGARGLSMVVADSTRRAXO 58
 DB 1 MAVALRFAAVARDPAPAAAPPRVGRBQVLAQAAAPRGRCRRRGLVVRGSGAAAVV 60
 QY 59 ADGGDGVLAGVLCGMEGLGDQLSVMKFGSSVSSAARNAEVAGLITPEERPVVY 118
 DB 61 LNKDD--AASVAAAAAASATG--FTVMKFGSSVSAARMEVADLLSPFEERPVV 115
 QY 119 LSAMGKTNNLLAGEKAVGCGVHVSEIEEMNVKSLIKITVDELGLP-XICNTSLVLEL 177
 DB 116 LSAMGKTNNLLAGEKAVGCGVHVSEIEEMNVKSLIKITVDELGLDLSIVSGLEEL 175
 QY 178 EOLLKGIAMMKELTPRTSDYLVSGECSTRIFSAYLNKIRVAKQYDAFDIGFITDDE 237
 DB 176 EOLLKGIAMMKELTPRTSDYLVSGECSTRIFAAYLNKIRVAKQYDAFDIGFITDDE 235
 QY 238 GNAIDLLEATYPAVAKRLHGMWIDPAIPVYTGFLGKWKSGAVTTLGRGSDLTATTIGK 297
 DB 236 TNADLLEATYPAVAKRLHGMWIDPAIPVYTGFLGKWKSGAVTTLGRGSDLTATTIGK 295
 QY 298 ALGLREIQVWKVDGVLTCDPNIYPAKTVPYLTPEEATLAYFGAQLVHPQMRPAREG 357
 DB 296 ALGLREIQVWKVDGVLTCDPNICANAIIPVYLTPEEATLAYFGAQLVHPQMRPAREG 355
 QY 358 DIPRVKSYNPKAPGTLTRQDMDGVLVTSIVLSKNTMLDIVSTRMLQGYGLAR 417
 DB 356 DIPRVKSYNPKAPGTLTRQDMDGVLVTSIVLSKNTMLDIVSTRMLQGYGLAR 413
 QY 418 VSGICYIEDLCISVDCVATSEVSVSLDPSKIMRELIQOASELDHVELEKIAIVRL 477
 DB 414 VSGI--FEDLGSVDCVATSEVSVSLDPSKIMCRELIQ--ELDHVELEKIAIVRL 469
 QY 478 LQORAIISLIGNEOSSILLEKTRVLRKSGVNVQMISQASKVMGLYHDSDAKALVE 537
 DB 470 LQHRISLISLIGNEOSSILLEKTRVLRKSGVNVQMISQASKVMGLYHDSDAKALVE 529
 QY 538 ALHQAFFEDVLSQVEAENL 557
 DB 530 ALHQAFFEDVLSQVEAENL 549

RESULT 3
 Q8RYL8 PRELIMINARY; PRT; 581 AA.
 ID Q8RYL8
 AC Q8RYL8;
 DT 01-JUN-2002 (TEMBLrel. 21, Created)
 DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
 DE Putative aspartate kinase (EC 2.7.2.4) (Aspartate_kinase).
 GN OSJNB0015N02.2.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriactoidae; Oryzae; Oryza.
 NC NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT Oryza sativa nipponbare (GA) genomic DNA, chromosome 1, BAC
RT clone:OSUNB0052012.1;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSPHO-L-
CC ASPARTATE.
CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.
DR EMBL: AF004330; BAB90744.1; -
DR Gmame; O89y18; -
DR GO: GO:0004072; F.aspartate kinase activity; IEA.
DR GO: GO:0016501; F.kinase activity; IEA.
DR GO: GO:0016740; F.transferase activity; IEA.
DR GO: GO:0008652; P.amino acid biosynthesis; IEA.
DR InterPro: IPR001048; Aa kinase.
DR InterPro: IPR001341; Aspartate_kinase.
DR Pfam: PF00696; aakinese; 1.
DR TIGRfam: TIGR00657; asp_kinase; 1.
DR PROSITE: PS00324; ASPARTOKINASE; 1.
KW Kinase; Transferase.
SQ SEQUENCE 581 AA; 62638 MW; AD8B0811EB906AAB CRC64;

Query Match 66.4%; Score 1868.5; DB 10; Length 581;
Best Local Similarity 66.0%; Pred. No. 2.9e-135;
Matches 409; Conservative 34; Mismatches 72; Indels 105; Gaps 11;

QY 1 MAIPVSAAP--RLVPSIIPASSG-----HVGGLACFTTGTGPRARG 43
DB 1 MAIALRLAALPLRLIVSPAPPAICAGDGRGGGGAARIGVLGRACRRRRGGGGLE 60
QY 44 LSNVVDNSR--RAKQADGGDVLGAPVLGHEGLDGLSVNKKFGSSSVSAAAPAE 102
DB 61 VAAAADDSRCRAKAV--GAAAAAETGGGIVGG--GQLSVNKKFGSSSVSAAARE 117
QY 103 VAGLLTFEERPVVLSAMGKTNNLLAGKAVGCGVIVHSEIEMNMYSLIKTV 162
DB 118 VAGLLTFEERPVVLSAMGKTNNLLAGKAVGCGVIVHSEIEMNMYSLIKTV 177
QY 163 ELGLP-XICNTSLYEELQKIAMKELTPTSDVLSFGCSMTRIFSAVINKIRYA 221
DB 178 ELALPSVHTMDLEQLKIAMKELTPTSDVLSFGCSMTRIFSAVINKIRYA 237
QY 222 RQYDADIGFITLDEFGNADILEATYPAVAKRLHGMIDDPALPVYTGFLGKMGSGAVT 281
DB 238 RQYDADIGFITLDEFGNADILEATYPAVAKRLHGMIDDPALPVYTGFLGKMGSGAVT 297
QY 282 TLGRGSDLTATTTGALGLREIQWVKVDVGLTCDPNITPRAKTVPIYLTPEATELAYF 341
DB 298 TLGRGSDLTATTTGALGLREIQWVKVDVGLTCDPNITPRAKTVPIYLTPEATELAYF 357
QY 342 GAQVLPQSRPARREGDIPRVKNSYNPAPGTLITRQDMXGVLVLSIYLSKNVMTL 401
DB 358 GAQVLPQSRPARREGDIPRVKNSYNPAPGTLITRQDMXGVLVLSIYLSKNVMTL 415
QY 402 DIVSTRMLGQYGLFARVSGICYEELICISVDCAVSESVSLDSKIMSELIQAS- 460
DB 416 DIVSTRMLGQYGLFARVSGICYEELICISVDCAVSESVSLDSKIMSELIQAS- 443
QY 461 -----ELDHVELEKIALVRLQORATIS 485
DB 444 PIYKQNVVITLQAGSDQSTFQVLFALPLPFGGLDHVELEKIALVRLQORATIS 503
QY 486 LIGNVQSSLILEKTVLRK-----SGVNVQMI SQASKVNSLIYHSDAKALVE 537
DB 504 LIGNVQSSLILEKTVLRK-----SGVNVQMI SQASKVNSLIYHSDAKALVE 537
QY 538 ALHQAFEDDVLQVEAENL 557
DB 559 ALHQAFEDDVLQVEAENL 578

ID 1D Q9XHC5 PRELIMINARY; PRT; 564 AA.
AC Q9XHC5;
DT 01-NOV-1999 (TRENBLREL. 12, Created)
DT 01-NOV-1999 (TRENBLREL. 12, Last sequence update)
DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
DE Precursor monofunctional aspartokinase (EC 2.7.2.4) (Aspartate
DE kinase).
OS Glycine max (soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_Taxid:3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Century;
RA Beau B.D., Frankard V., Jacobs M., Matthews B.F.;
RT "Isolation and characterization of a cDNA clone encoding a
RT monofunctional aspartokinase."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSPHO-L-
CC ASPARTATE.
CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.
DR EMBL: AF135862; AAD41796.1; -
DR GO: GO:0016597; P.amino acid binding; IEA.
DR GO: GO:0004072; F.aspartate kinase activity; IEA.
DR GO: GO:0016301; F.kinase activity; IEA.
DR GO: GO:0016740; F.transferase activity; IEA.
DR GO: GO:0008652; P.amino acid biosynthesis; IEA.
DR GO: GO:0008152; P.metabolism; IEA.
DR InterPro: IPR001048; Aa kinase.
DR InterPro: IPR002912; ACT.
DR InterPro: IPR001341; Aspartate_kinase.
DR Pfam: PF00696; aakinese; 1.
DR TIGRfam: TIGR00657; asp_kinase; 1.
DR PROSITE: PS00324; ASPARTOKINASE; 1.
KW Kinase; Transferase.
SQ SEQUENCE 564 AA; 61285 MW; 5835C4F855FC6157 CRC64;

Query Match 65.2%; Score 1835; DB 10; Length 564;
Best Local Similarity 70.6%; Pred. No. 1.1e-132;
Matches 389; Conservative 42; Mismatches 80; Indels 40; Gaps 7;

QY 37 GPGGARGLGVVADSTRRAKQADGDEVLGAPVL----- 71
DB 10 GVGGLAVNSVRSRLHCKSQI--GPAALGAPVCARRRWGNRAVSYTTCKASTDYLE 67
QY 72 -----GGGMEGLDGLSVNKKFGSSSVSAAAPAEVAGLTFEERPVVLSAMGKT 125
DB 68 KNATENGWVSSSG-ETSFYCVNKKFGSSSVASADRMKEVATLISFPEERPIVLSAMGKT 126
QY 126 TNNLLAGKAVCGVIVHSEIEMNMYSLIKTVDELGLP-XICNTSLYEELQKGI 184
DB 127 TNNLLAGKAVCGVIVHSEIEMNMYSLIKTVDELGLP-XICNTSLYEELQKGI 186
QY 185 AMKELTPTSDVLSFGCSMTRIFSAVINKIRYKQAQYAPDGLFTTPEPGNADILE 244
DB 187 AMKELTPTSDVLSFGCSMTRIFSAVINKIRYKQAQYAPDGLFTTPEPGNADILE 246
QY 245 ATYPAVAKRLHGMIDDPALPVYTGFLGKMGSGAVTTLGRGSDLTATTTGALGLREI 304
DB 247 ATYPAVAKRLHGMIDDPALPVYTGFLGKMGSGAVTTLGRGSDLTATTTGALGLREI 306
QY 305 QWVKVDVGLTCDPNITPRAKTVPIYLTPEATELAYFQAQVLPQSRPARREGDIPRVK 364
DB 307 QWVKVDVGLTCDPNITPRAKTVPIYLTPEATELAYFQAQVLPQSRPARREGDIPRVK 366
QY 365 NSYNPAPGTLITRQDMXGVLVLSIYLSKNVMTLIVSTRMLGQYGLFARVSGICYI 424
DB 367 NSYNPAPGTLITRQDMXGVLVLSIYLSKNVMTLIVSTRMLGQYGLFARVSGICYI 422
QY 425 EDLICISVDCAVSESVSLDSKIMSELIQASSELDRHVELEKIALVRLQORATIS 484

DB 423 EELGTSVDVATSEVSVSLTLPDKLMSRELIIQASELDHYVELEKIAVNNLQNRSLI 482
 QY 485 SLIGNVEQSLILEKTRGYLRKSGVNVQMSQASKNMSLITVDSKALVEALHQAFF 544
 DB 483 SLIGNVRSILIERLSRLTGLTVQVMSQASKNMSLVNDSEAEQCRALHSAFF 542
 QY 545 EDDVLSQVEAE 555
 DB 543 ESE-LSLELMD 552
 RESULT 5
 Q9LYU8 PRELIMINARY; PRT; 569 AA.
 ID Q9LYU8
 AC Q9LYU8:
 DT 01-OCT-2000 (TEMBLrel. 15, Created)
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Aspartate kinase (EC 2.7.2.4) (AT5G13280/T31B5_100)
 DE (Aspartokinese).
 GN T31B5.100.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC euroside II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
 RA Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai U., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Becker J.R.;
 RT Arabidopsids cDNA clones";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Shim P., Bowser L., Carninci P.,
 RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
 RA Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai U., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
 RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
 RA Southwick A., Tang C.C., Toriumi M., Wallender B.K., Wong C., Wu H.C.,
 RA Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,
 RA Becker J.R.;
 RT "Arabidopsids ORF clones";
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSPHO-L-
 CC ASPARTATE.
 CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.
 DR EMBL: AL163491; CAB86635.1; -
 DR EMBL: AY057674; AAL15305.1; -
 DR EMBL: BT000493; AAN18062.1; -
 DR PIR: T48575; T48575.
 DR GO: GO:0016597; P.amino acid binding; IEA.
 DR GO: GO:0004072; F.aspartate kinase activity; IEA.
 DR GO: GO:0016501; F.kinase activity; IEA.
 DR GO: GO:0016740; F.transferase activity; IEA.
 DR GO: GO:0008652; P.amino acid biosynthesis; IEA.
 DR GO: GO:0008152; P.metabolism; IEA.
 DR InterPro: IPR001046; Aa_kinase.

DR InterPro: IPR002912; ACT.
 DR InterPro: IPR001341; Aspartate_kinase.
 DR Pfam: PF00696; aak_kinase; 1.
 DR TIGRfam: TIGR00657; asp_kinases; 1.
 DR PROSITE: PS00324; ASPARTOKINASE; 1.
 KW kinase; transferase.
 SQ SEQUENCE 569 AA; 62298 MW; F66A3F4E84DC429 CRC64;
 Query Match 63.6%; Score 1789; DB 10; Length 569;
 Best Local Similarity 71.8%; Pred. No. 3.8e-129;
 Matches 369; Conservative 54; Mismatch 79; Indels 12; Gaps 6;
 QY 44 LSMVADSTRRAQADGDCGVLGAPYLGLGMEGLD---QLSVMKFGSSSVSAAR 99
 DB 44 LSLPFGDSSIRKYSGSGSRIVRA-VLEEKTEALTEVDEKIGTCVMKFGSSVSAAR 102
 QY 100 MAEVAGLLTTPPERPVVYVSAMGKTNNLLGKRAVGGCVHVEIEEMNVKSLHX 159
 DB 103 MKEVADLITPPESPVIVSAMGKTNNLLGKRAVGGCVHVEIEELSTIKELHR 162
 QY 160 TVDELGL-PIXCNSTLYELEQLKGIAMKELTPRTSDYLVSPGECMSIRIFSAVUNKR 218
 DB 163 TKELNIDPSVILVYLEELQGLKGIAMKELTRTDYLVSPGECLSIRIFAYLNTIG 222
 QY 219 VKARQYDAFDIGFTTTDFGNAQDILETYPAVAKRLHGMVQDPALPVYTGFGKWKSG 278
 DB 223 VKARQYDAFEIGFTTTDFGNDILETYPAVAKRLVDDMDHPAPVIVGFGKWKSG 282
 QY 279 AVTTGGRGSDLTATTTGKALGRLRIQVMDVDGVLTCNDINYPHAKTVPYLFEETEL 338
 DB 283 AVTTGGRGSDLTATTTGKALGRLRIQVMDVDGVLTCNDITPKRTPPYLTFDDAAL 342
 QY 339 AVFGAQLVHPQSKRPAREGDIPIVAKSYNPKAPGTLITQRMDXGLVLSIVKSNV 398
 DB 343 AVFGAQLVHPQSKRPAREGDIPIVAKSYNPKAPGTLITQRMTS--ILTSIVLKRNV 400
 QY 399 TMLDVSITRMGGVGLARVSGICIEDLCISVDCVATSEVSVSLTDPKMSRELIIQ 458
 DB 401 TMLDIASITRMGGVGLAKVRSI--FELGISVDVATSEVSVSLTDPKMSRELIIQ 458
 QY 459 ASELDHYVELEKIAVRLIQORAIISLIGNVQSSSILEKTRGYLRKSGVNVQMSQGA 518
 DB 459 --ELDHYVELEKIAVNNLKGRAIISLIGNVQSSSILIERAPHVYTKGVNVQMSQGA 516
 QY 519 SKVNSLIVHDSDAKALVEALHQAFFEDVLSQV 552
 DB 517 SKVNISITVNEAEAGCVQLHKSFFESGDLSEL 550
 RESULT 6
 O23152 PRELIMINARY; PRT; 569 AA.
 ID O23152
 AC O23152:
 DT 01-JAN-1998 (TEMBLrel. 05, Created)
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Aspartate kinase precursor (EC 2.7.2.4) (Aspartokinese).
 GN AK-LYS1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC euroside II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97351557; PubMed=9207839;
 RA Frankard V., Vaucherin M., Jacobs M.;
 RT "Molecular characterisation of an Arabidopsis thaliana cDNA coding for
 a monofunctional aspartate kinase";
 RL Plant Mol. Biol. 34:233-242 (1997).
 CC -1- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSPHO-L-
 CC ASPARTATE.

CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.
 DR EMBL; X98873; CA67376.1; -;
 DR GO; GO:0016597; F:amino acid binding; IEA.
 DR GO; GO:0004072; F:aspartate kinase activity; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR001048; Aa_kinase.
 DR InterPro; IPR002912; ACT.
 DR InterPro; IPR001341; Aspartate_kinase.
 DR Pfam; PF00696; aak_kinase; 1.
 DR Pfam; PF01842; ACT; 2.
 DR TIGRFAMs; TIGR00657; asp_kinases; 1.
 DR PROSITE; PS00324; ASPARTOKINASE; 1.
 KM Kinase; Signal; Transferase.
 FT SIGNAL 1 86
 FT CHAIN 87 569 ASPARTATE KINASE.
 SQ SEQUENCE 569 AA; 62272 MW; CD5323120D94DD3 CRC64;
 Query Match 63.4%; Score 1784; DB 10; Length 569;
 Best Local Similarity 71.6%; Pred. No. 9, 2e-129;
 Matches 368; Conservative 55; Mismatches 79; Indels 12; Gaps 6;
 QY 44 LSNVVDSTRRKQADGQGVIGAPVIGLGMESGLD---QLSVMKFGSSVSAAAR 99
 DB 44 LSLPTGDSIRKVSQSGSNIVRA-VLEKKTETATEDEKGITCMKFGSSVSAAR 102
 QY 100 MAEYAGLITFPERPVPVVLASAKTNNLLAGEKAVGCVIHVSEIEMNVKSLHX 159
 DB 103 MKEVADILTFPEESPVTIVASAKTNNLLAGEKAVGCVINASEIELSLIKELHR 162
 QY 160 TVDEELG-EXICNTSLYELEQLKGTAMKELTPRSDIVYSGEOMRISAYINKR 218
 DB 163 TVKELNTDPSVILTYLEELQLKGTAMKELTRDYVSGECLSTRIPAAVYNTIG 222
 QY 219 VKARQYDAFIDGITTFDEGNADILEATYPAVAKRLHGMIDPAIPVVTGFLGKQMSG 278
 DB 223 VKARQYDAFELGITTFDDFTNGDILEATYPAVAKRLYDMDMDPAIPVITGFLGKQMSG 282
 QY 279 AVTTLGRGSDLTATTIGKALGFEIOWKDVGVITCDPNYIPAKYIPYITPEAEEL 338
 DB 283 AVTTLGRGSDLTATTIGKALGFEIOWKDVGVITCDPTIKRATPPVYITPEAEEL 342
 QY 339 AYFGAQLHPQSMRPAEEDIPVRKNSYNPKAPGTLITRQMDGLVLTSLYKSNV 398
 DB 343 AYFGAQLHPQSMRPAEEDIPVRKNSYNPKAPGTLITRQMDGLVLTSLYKSNV 400
 QY 399 TMDIVSTRMLGQYGLFARVSGICYIEDLCISVDCVATSEVSVSLDPSKIMSRLEIQ 458
 DB 401 TMDIVSTRMLGQYGLFARVSGICYIEDLCISVDCVATSEVSVSLDPSKIMSRLEIQ 458
 QY 459 ASELDHYVEELKAIATRLIQARITSLIGNVOSSLILEKGRVARKSNVNMVMSOGA 518
 DB 459 ASELDHYVEELKAIATRLIQARITSLIGNVOSSLILEKGRVARKSNVNMVMSOGA 516
 QY 519 SKVMSLIIVHSDAKALVEALHQAFFEDVLSQV 552
 DB 517 SKVMSLIIVHSDAKALVEALHQAFFEDVLSQV 550
 RESULT 7
 Q9FM04 PRELIMINARY; PRT; 544 AA.
 AC Q9FM04;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Lysine-sensitive aspartate kinase (EC 2.7.2.4)
 DE (Aspartokinase)
 OS Arabidopsis thaliana (Mouse-ear cress)
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

CC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 CX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RX MEDLINE=98162728; PubMed=9501997;
 RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
 Tabata S.;
 RA "Structural analysis of Arabidopsis thaliana chromosome 5. III.
 RT Sequence features of the regions of 1,191,918 bp covered by seventeen
 RT physically assigned P1 clones."
 RI DNA Res. 4:401-414 (1997).
 CC -1- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSHO-L-
 CC ASPARTATE.
 CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.
 DR EMBL; AB007650; BAB08285.1; -;
 DR GO; GO:0016597; F:amino acid binding; IEA.
 DR GO; GO:0004072; F:aspartate kinase activity; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR001048; Aa_kinase.
 DR InterPro; IPR002912; ACT.
 DR InterPro; IPR001341; Aspartate_kinase.
 DR Pfam; PF00696; aak_kinase; 1.
 DR Pfam; PF01842; ACT; 2.
 DR TIGRFAMs; TIGR00657; asp_kinases; 1.
 DR PROSITE; PS00324; ASPARTOKINASE; 1.
 KM Kinase; Transferase.
 SQ SEQUENCE 544 AA; 59604 MW; 7DBCFDC1138645AC CRC64;
 Query Match 62.1%; Score 1747.5; DB 10; Length 544;
 Best Local Similarity 74.5%; Pred. No. 5, 6e-126;
 Matches 351; Conservative 50; Mismatches 65; Indels 5; Gaps 3;
 QY 78 GLGDOLSVYMKFGSSVSAAAEVAGLITFPERPVPVVLASAKTNNLLAGEKAV 137
 DB 77 GTGKELTCVMKFGSSVSABRKEVANIILSPDEREIVIVASAKTNNLLAGEKAV 136
 QY 138 GCGVIVHSEIEMNVKSLHITVDELG-EXICNTSLYELEQLKGTAMKELTPRSD 196
 DB 137 TCGVTVNESIEELSLFKELHTRAEHLETTVIEKHLEGLHQLKGTAMKELTRD 196
 QY 197 YLVSPGECSTRIFSAVYLNKIRKARQYDAFIDGITTFDEGNADILEATYPAVAKRLHG 256
 DB 197 YLVSPGECSTRIFSAVYLNKIRKARQYDAFIDGITTFDDFTNGDILEATYPAVAKRLHG 256
 QY 257 DWIQDPAIPVVTGFLGKMSGAVTTLGRGSDLTATTIGKALGFEIOWKDVGVITC 316
 DB 257 DWKRENAVPPVTVYLOKGRSCAITTLGRGSDLTATTIGKALGFEIOWKDVGVITC 316
 QY 317 DPNYIPAKYIPYITPEAEELAYFGAQLHPQSMRPAEEDIPVRKNSYNPKAPGTLI 376
 DB 317 DPNYIPAKYIPYITPEAEELAYFGAQLHPQSMRPAEEDIPVRKNSYNPKAPGTLI 376
 QY 377 TTRQMDXGLVLTSLYKSNVMTMDIVSTRMLGQYGLFARVSGICYIEDLCISVDCVAT 436
 DB 377 TTRQMDXGLVLTSLYKSNVMTMDIVSTRMLGQYGLFARVSGICYIEDLCISVDCVAT 432
 QY 437 SEVSVSLDPSKIMSRLEIQASELDHYVEELKAIATRLIQARITSLIGNVOSSLI 496
 DB 437 SEVSVSLDPSKIMSRLEIQASELDHYVEELKAIATRLIQARITSLIGNVOSSLI 492
 QY 497 LEKTRVLRKSGVNVQMISQASKVMSLIIVHSDAKALVEALHQAFFEDD 547
 DB 493 LEKTRVLRKSGVNVQMISQASKVMSLIIVHSDAKALVEALHQAFFEDD 543
 RESULT 8
 O23653 PRELIMINARY; PRT; 544 AA.
 ID O23653
 AC O23653;

DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
 DE Lysine-sensitive aspartate kinase (EC 2.7.2.4)
 DE (Aspartokinase).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsi.
 NCBI Taxid=3702;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9731562; PubMed=9207844;
 RA Teng G., Zhu-Shimoni J.X., Amir R., Zehori I.B., Galili G.;
 RT "Cloning and expression of an Arabidopsis thaliana cDNA encoding a
 RT monofunctional aspartate kinase homologous to the lysine-sensitive
 RT enzyme of Escherichia coli.";
 RL Plant Mol. Biol. 34:287-294(1997).
 CC -1- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSPHO-L-
 CC ASPARTATE.
 CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.
 DR EMBL: U62020; AAB63104.1; -;
 DR GO: GO:0016597; F:amino acid binding; IEA.
 DR GO: GO:0004072; F:aspartate kinase activity; IEA.
 DR GO: GO:0016301; F:kinase activity; IEA.
 DR GO: GO:0016740; F:transferase activity; IEA.
 DR GO: GO:000652; P:amino acid biosynthesis; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR001048; Aa_kinase.
 DR InterPro: IPR002912; ACT.
 DR InterPro: IPR001341; Aspartate_kinase.
 DR Pfam: PF00696; aak_kinase; 1.
 DR Pfam: PF01842; ACT; 2.
 DR TIGRFAMs: TIGR00657; asp_kinases; 1.
 DR PROSITE: PS00324; ASPARTOKINASE; 1.
 DR Kinase; Transferase.
 KM KINASE 544 AA; 59506 MW; 2C204BCEB23DD80 CRC64;
 SQ SEQUENCE

Query Match 61.8%; Score 1737.5; DB 10; Length 544;
 Best Local Similarity 74.1%; Pred. No. 3.3e-125;
 Matches 349; Conservative 50; Mismatches 67; Indels 5; Gaps 3;

QY 78 GLGDQSVYMKFGSSVSSAARMAEYAGLILTPPEERPVVLSAMGKTTNNLLAGEKAV 137
 DB 77 GTGKELTCVMKFGSSVSSAARKEVANLIFSPDERPVIVLSAMGKTTNNLLXAGEKAV 136
 QY 138 GCGVIVHSEIEEMNNVKSLLIKTVDELGL-PXICNTSLYELEQLKGIAMMKELTPRTSD 196
 DB 137 TCGVTVGSEIEELSLFKELHLRTHAHLGVETTVIEKLEGLHQLKGISMKKELTLRTRD 196
 QY 197 YLVSGEGCSTRIFSAVYLNKIRVAKQYDAFDIGFTTDFEGNADILEATYPAVAKLHG 256
 DB 197 YLVSGEGCSTRIFSAVYLNKIRVAKQYDAFDIGFTTDFEGNADILEATYPAVAKLHG 256
 QY 257 DNIQDPAIPVVTGFLGKMGSGAVTTLGRGSDLTATTIGKALGRIEIQWKVDVGLTLC 316
 DB 257 DMSKENAVPVYTGVLGKMRSCAITTLGRGSDLTATTIGKALGRIEIQWKVDVGLTLC 316
 QY 317 DENIYPHAKTVPYLTFEATELAFGAQVLFHPQMRPARBEDIIVRVKSNINPAFGTLI 376
 DB 317 DPAIYPGAOSVPLTFDEAAELAFGAQVLFHPQMRPARBEDIIVRVKSNINPAFGTLI 376
 QY 377 TRORDMXGLVLTSLVLSKSNVTMLDIVSTMLGQYGLFARVSGICYEDLCISVDVAT 436
 DB 377 TRSRDMSK-AVLTSLVLSKSNVTMLDIVSTMLGQYGLFARVSGICYEDLCISVDVAT 436
 QY 437 SEVSVSLSDPKIMSRRELIOQASGLDHVVELEKIAIVRLIOQRATISLIGNVQSSLI 496
 DB 433 SEVSISLTPDPAKLMRELIOQRVNELDNLVEELEKIAVYKLIQRRSIIISLIGNVQSSLI 492
 QY 497 LEVTGVLKRSQVNVOMISQAGKVMSLIVSDPAKLLVEALHQAFFEDD 547
 DB 493 LEKVFQVFRSNGVNVOMISQAGKVMSLIVSDPAKLLVEALHQAFFEDD 543

RESULT 9
 ID Q9FY44 PRELIMINARY; PRT; 544 AA.
 AC Q9FY44;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
 DE Aspartate kinase precursor (EC 2.7.2.4) (Aspartokinase).
 GN AK-LYS2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsi.
 NCBI Taxid=3702;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia; TISSUE=Leaf;
 RA Frankard V.M.S., Vauterin M., Jacobs M.;
 RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSPHO-L-
 CC ASPARTATE.
 CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.
 DR EMBL: Y16255; CAC06395.1; -;
 DR GO: GO:0016597; F:amino acid binding; IEA.
 DR GO: GO:0004072; F:aspartate kinase activity; IEA.
 DR GO: GO:0016301; F:kinase activity; IEA.
 DR GO: GO:0016740; F:transferase activity; IEA.
 DR GO: GO:000652; P:amino acid biosynthesis; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR001048; Aa_kinase.
 DR InterPro: IPR002912; ACT.
 DR InterPro: IPR001341; Aspartate_kinase.
 DR Pfam: PF00696; aak_kinase; 1.
 DR Pfam: PF01842; ACT; 2.
 DR TIGRFAMs: TIGR00657; asp_kinases; 1.
 DR PROSITE: PS00324; ASPARTOKINASE; 1.
 KM Kinase; Signal; Transferase.
 FT SIGNAL 1 84
 SQ SEQUENCE 544 AA; 59578 MW; 3308371224C309A6 CRC64;

Query Match 61.8%; Score 1737.5; DB 10; Length 544;
 Best Local Similarity 74.1%; Pred. No. 3.3e-125;
 Matches 349; Conservative 50; Mismatches 67; Indels 5; Gaps 3;

QY 78 GLGDQSVYMKFGSSVSSAARMAEYAGLILTPPEERPVVLSAMGKTTNNLLAGEKAV 137
 DB 77 GTGKELTCVMKFGSSVSSAARKEVANLIFSPDERPVIVLSAMGKTTNNLLXAGEKAV 136
 QY 138 GCGVIVHSEIEEMNNVKSLLIKTVDELGL-PXICNTSLYELEQLKGIAMMKELTPRTSD 196
 DB 137 TCGVTVGSEIEELSLFKELHLRTHAHLGVETTVIEKLEGLHQLKGISMKKELTLRTRD 196
 QY 197 YLVSGEGCSTRIFSAVYLNKIRVAKQYDAFDIGFTTDFEGNADILEATYPAVAKLHG 256
 DB 197 YLVSGEGCSTRIFSAVYLNKIRVAKQYDAFDIGFTTDFEGNADILEATYPAVAKLHG 256
 QY 257 DNIQDPAIPVVTGFLGKMGSGAVTTLGRGSDLTATTIGKALGRIEIQWKVDVGLTLC 316
 DB 257 DMSKENAVPVYTGVLGKMRSCAITTLGRGSDLTATTIGKALGRIEIQWKVDVGLTLC 316
 QY 317 DENIYPHAKTVPYLTFEATELAFGAQVLFHPQMRPARBEDIIVRVKSNINPAFGTLI 376
 DB 317 DPAIYPGAOSVPLTFDEAAELAFGAQVLFHPQMRPARBEDIIVRVKSNINPAFGTLI 376
 QY 377 TRORDMXGLVLTSLVLSKSNVTMLDIVSTMLGQYGLFARVSGICYEDLCISVDVAT 436
 DB 377 TRSRDMSK-AVLTSLVLSKSNVTMLDIVSTMLGQYGLFARVSGICYEDLCISVDVAT 436
 QY 437 SEVSVSLSDPKIMSRRELIOQASGLDHVVELEKIAIVRLIOQRATISLIGNVQSSLI 496
 DB 433 SEVSISLTPDPAKLMRELIOQRVNELDNLVEELEKIAVYKLIQRRSIIISLIGNVQSSLI 492

QY 497 LEKTRVLRKSGVNVQMSIGASKVNLSLIVHSDAKALVEALHQAFFEDD 547
DB 493 LEKTRVLRKSGVNVQMSIGASKVNLSLIVHSDAKALVEALHQAFFEDD 543

RESULT 10

Q9S702 PRELIMINARY; PRT; 559 AA.
AC Q9S702;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Putative aspartate kinase (EC 2.7.2.4) (Aspartokinase).
GN P109.20 OR F28J7.35.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
CC NCBI_TaxID=3702;

QY SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Bentto M., Creasy T.H., Haas B.,
RA Rouning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC P109 genomic sequence."
RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volfovsky N., Town C.D., Troughan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation." Genome Biol. 0:0-0(2002).
RL [3]
RN SEQUENCE FROM N.A.
RA Brever V., Troughan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana."
RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSPHO-L-
CC ASPARTATE.

CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.
CC EMBL; AC011664; AAF14833.1; -;
DR EMBL; AC010797; AAF03452.1; -;
DR EMBL; AY088366; AAM65905.1; -;
DR GO; GO:0016597; F:amino acid binding; IEA.
DR GO; GO:0004072; F:aspartate kinase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001048; Aa_kinase.
DR InterPro; IPR002912; ACT.
DR InterPro; IPR001341; Aspartate_kinase.
DR InterPro; IPR00696; aak_kinase; 1.
DR Pfam; PF01842; ACT; 2.
DR TIGRPFAM; TIGR00657; asp_kinases; 1.
DR PROSITE; PS00324; ASPARTOKINASE; 1.
KW Kinase; Transferase.
SQ SEQUENCE 559 AA; 61215 MW; FRC684A0B814F349 CRC64;

Query Match 59.6%; Score 1675.5; DB 10; Length 559;
Best Local Similarity 73.5%; Pred. No. 2,1e-120;
Matches 347; Conservative 47; Mismatches 71; Indels 7; Gaps 4;

QY 82 QLSVVRKFGSSVSAAARMAEVLITPPEERPVVVLASMGKTNNLLAGKAVGCCV 141
DB 82 KLTCTMKFGSSVSAAARMAEVLITPPEERPVVVLASMGKTNNLLAGKAVGCCV 141
QY 142 IHVSEIEWNNVKSLLIKYVDELGL-EXICNTSLVELEQLKGIAMKKEITPTSTYLV 200

DB 142 TNVDITIEELSYIEHIRTAEHCFTAVIAEHLEGLQKGVAMKEILTISRDIYS 201
QY 201 FGECMSTRIRFSAYLNIRVKARQYDAFDIGITTTDEFGNADILEATYPAVAKRLHGM 260
DB 202 FGECMSTRIRFSAYLNIRVKARQYDAFDIGITTTDEFGNADILEATYPAVAKRLHGM 261
QY 261 DPALPVYTGFLGKSGKSGAVTTLGRGSDLTATTIGKALGREIOWMDVGVCLCDPNI 320
DB 262 ENALPVYTGFLGKSGKSGAVTTLGRGSDLTATTIGKALGREIOWMDVGVCLCDPNI 321
QY 321 YPAKTVPYLTPEERATELAVFGAOVLPQSMRPAEGDIPVKNKSNYPKAPGTLITROR 380
DB 322 YCAQAPPHITFDEAEELAVFGAOVLPQSMRPAEGDIPVKNKSNYPKAPGTLITROR 381
QY 381 DMGXGLVLTSTYLRKSNVMTDIVSRMLGQGFILARVSGITIIDLCISDCAVTSFVS 440
DB 382 DMKR-AVLISYLRKSNVMTDIVSRMLGQGFILARVSGITIIDLCISDCAVTSFVS 437
QY 441 VSYSLDPSKIMSELIQOASELDHVEELEKTAIVRLQOQRAIISLIGNEQSSILIBKT 500
DB 438 ISLTDPKSCSELIQH--ELDQVEELEKTAIVNLRHRSITSLIGNVQSSITLIEKG 495
QY 501 GRVLRKSGVNVQMSIGASKVNLSLIVHSDAKALVEALHQAFFEDDVLQY 552
DB 496 GRVLRKSGVNVQMSIGASKVNLSLIVHSDAKALVEALHQAFFEDDVLQY 547

RESULT 11

Q9MAX0 PRELIMINARY; PRT; 354 AA.
AC Q9MAX0;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Aspartate kinase (EC 2.7.2.4) (Aspartokinase) (Fragment).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
CC NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Kiyota S.;
RT "Lysine sensitive aspartate kinase from rice."
RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-ASPARTATE = ADP + 4-PHOSPHO-L-
CC ASPARTATE.

CC -1- SIMILARITY: BELONGS TO THE ASPARTOKINASE FAMILY.
CC EMBL; AB042521; BA95630.1; -;
DR Gramene; Q9MAX0; F:amino acid binding; IEA.
DR GO; GO:0016597; F:amino acid binding; IEA.
DR GO; GO:0004072; F:aspartate kinase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001048; Aa_kinase.
DR InterPro; IPR002912; ACT.
DR InterPro; IPR001341; Aspartate_kinase.
DR InterPro; IPR00696; aak_kinase; 1.
DR Pfam; PF01842; ACT; 2.
DR TIGRPFAM; TIGR00657; asp_kinases; 1.
KW Kinase; Transferase.
FT NON_TER 1
SQ SEQUENCE 354 AA; 38885 MW; 88B668815627A040 CRC64;

Query Match 55.8%; Score 1570; DB 10; Length 354;
Best Local Similarity 88.2%; Pred. No. 1,4e-112;
Matches 315; Conservative 18; Mismatches 18; Indels 5; Gaps 3;

QY 201 FGECMSTRIRFSAYLNIRVKARQYDAFDIGITTTDEFGNADILEATYPAVAKRLHGM 260
DB 1 FGECMSTRIRFSAYLNIRVKARQYDAFDIGITTTDEFGNADILEATYPAVAKRLHGM 260

QY 261 DPAPVVTGFLGKMGKSGAVTTLGRGSDLTATTIGALGLREIQVMKDVDTGTCDPNI 320
 DB 61 DPAPVVTGFLGKMGKSGAVTTLGRGSDLTATTIGALGLREIQVMKDVDTGTCDPNI 120
 QY 321 YPAKTYPTVLTPEARELAVFGAQLVHPQSMRPARBGDI PVRVKSYPKAPGTLITROR 360
 DB 121 YPAATVPVLTPEARELAVFGAQLVHPQSMRPARBGDI PVRVKSYPKAPGTLITROR 180
 QY 381 DMDXGLVLTSLVLSKNTWLDIVSTRMLGQYGFARVSGICIEDLCISVDCVATSEVS 440
 DB 181 EMOK--VLTSLVLSKNTWLDIVSTRMLGQYGFARVSGICIEDLCISVDCVATSEVS 236
 QY 441 VVSITDPSKTSWRELLIQASSELDPHYVEELEKTAIYVLLQQRALISLIGNVQSSLIETKT 500
 DB 237 VVSITDPSKTSWRELLIQ--ELDPHYVEELEKTAIYVLLQQRALISLIGNVQSSLIETKT 294
 QY 501 GRVLRKSGVNVQMIISQASKVNMSLIVHSDAKALVHALQAFFEDVLSQVEAENTL 557
 DB 295 FGVLRSKGVNVQMIISQASKVNMSLIVHSDAKALVHALQAFFEDVLSQVEAENTL 351
 RESULT 12
 Q8KG73 PRELIMINARY; PRT; 470 AA.
 ID Q8KG73
 AC Q8KG73;
 DT 01-OCT-2002 (TEMBLrel. 22, Created)
 DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Aspartokinase.
 GN LYSC OR CT0095.
 OS Chlorobium tepidum.
 OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
 OC Chlorobium.
 NCBI_TaxID=1097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TLS / ATCC 49652 / DSM 12025;
 RX MEDLINE=22103685; PubMed=1203901;
 RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
 RA Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
 RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
 RA Holt I., Umayar L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
 RA Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
 RA Vamathevan J., Khorai H., Bryant D.A., Fraser C.M.,
 RA "The complete genome sequence of Chlorobium tepidum TLS, a
 RT photosynthetic, anaerobic, green-sulfur bacterium",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
 DR EMBL; AE012788; AAM71343.1;
 DR TIGR; CT0095;
 DR GO; GO:0016597; F:amino acid binding; IEA.
 DR GO; GO:0004072; F:aspartate kinase activity; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001046; Aa_kinase.
 DR InterPro; IPR002912; ACT.
 DR InterPro; IPR001341; Aspartate_kinase.
 DR InterPro; IPR000847; HTH_LysR.
 DR Pfam; PF00696; aakinese; 1.
 DR Pfam; PF01842; ACT; 2.
 DR TIGRFAMs; TIGR00657; asp_kinase; 1.
 DR PROSITE; PS00324; ASPARTOKINASE; 1.
 DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
 DR Kinase; Complete proteome.
 QW SEQUENCE 470 AA; 50416 MW; DD6FE74EFA2730C5 CRC64;
 Query March 24.7%; Score 695.5; DB 16; Length 470;
 Best Local Similarity 37.6%; Pred. No. 5.9e-45;
 Matches 164; Conservative 93; Mismatches 164; Indels 49; Gaps 16;

QY 85 VYMKRGSSVSSAAMAEVAGLILTFPER---PVVYLSAMKTTNNLLAGEKAVGCC 140
 DB 2 VYMKRGSSVSSAAMAEVAGLILTFPER---PVVYLSAMKTTNNLLAGEKAVGCC 57
 QY 141 VTH-----VSEIEENMKSLIKTVDLGLPKICNTSLY--ELEQLKGIAMKELTPR 193
 DB 58 CLEBAQVGEVROCHDLIGELISEBLOQEVIAKTEVLTRELRTEGIEIVGELTER 117
 QY 194 TSDYLVSGECNSTRIFSAVYLNKIRYKARQVADIDGFIITDFG---NADILKATPA 249
 DB 118 SKDRFCSFEGELISTSVFAALNEAGVSCMKIDVATV-MITDDRGRFARPLAICQKNTSE 176
 QY 250 VAKRLHGMIDQPPAPVVTGFLGKMGKSGAVTTLGRGSDLTATTIGALGLREIQVMK 309
 DB 177 IIRKPL-----LDAGTVVTVQYIG-ALEGRTTIGRGSDLSALPFAMHSESIEIWD 231
 QY 310 VDGLTCDENIYPAKTYPTVLTPEARELAVFGAQLVHPQSMRPARBGDI PVRVKSYP 369
 DB 232 VDGLTCDENIYPAKTYPTVLTPEARELAVFGAQLVHPQSMRPARBGDI PVRVKSYP 231
 QY 370 KAPGTLITRORMDXGLV---VLTSLVLSKNTWLDIVSTRMLGQYGFARVSGICIED 426
 DB 292 DSKGTLITRORMDXGLV---VLTSLVLSKNTWLDIVSTRMLGQYGFARVSGICIED 349
 QY 427 LCISVDCVATSEVSVSILDPKTSWRELLIQASSELDPHYVEELEKTAIYVLLQQRALISL 486
 DB 350 FGISVEMISTSEVSISLVDDAVV-SEPLIALGALGCV--EIE-----HKVATVSV 358
 QY 487 IG-NVEQSSLIETKTGRVLRK--SGVNVQMIISQASKVNMSLIVHSDAKALVHALQAFF 544
 DB 399 VGDNLKMSKGV--AGRIPLSLRNVMIMISQASEINVGVDSDVQAIVSLHCEFF 455
 QY 545 EDDVLSQVEA 554
 DB 456 AE---SQGDA 462
 RESULT 13
 Q8T230 PRELIMINARY; PRT; 467 AA.
 ID Q8T230
 AC Q8T230;
 DT 01-JUN-2002 (TEMBLrel. 21, Created)
 DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Aspartokinase.
 GN LYSC OR MK0109.
 OS Methanopyrus kandleri.
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
 OC Methanopyrus.
 NCBI_TaxID=3320;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AV19 / DSM 6324 / JCM 9639;
 RX MEDLINE=21927647; PubMed=11930014;
 RA Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,
 RA Scharberlin O.V., Shakhova V.V., Belova G.I., Arvid L.,
 RA Metale P.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
 RA Malyn A.G., Koonin E.V., Kozaykin S.A., Kozaykin S.A.,
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
 RT and monophyly of archaeal methanogens.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
 DR EMBL; AE010310; AAM01326.1;
 DR GO; GO:0016597; F:amino acid binding; IEA.
 DR GO; GO:0004072; F:aspartate kinase activity; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR001048; Aa_kinase.
 DR InterPro; IPR002912; ACT.
 DR InterPro; IPR001341; Aspartate_kinase.
 DR InterPro; IPR005260; Asp_kin_monoph.
 DR Pfam; PF00656; aakinese; 1.

Search completed: March 23, 2004, 13:17:34
Job time : 77 secs

```
RN SEQUENCE FROM N.A.
RP STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Ancoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McGwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.B., Grathame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.U., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanobrevibacterium smithii reveals extensive metabolic
RT and physiological diversity.";
RI Genome Res 12:532-542(2002).
DR EMBL; AE010670; AM03585.1;
DR GO; GO:0016597; F:amino acid binding; IEA.
DR GO; GO:0004072; F:aspartate kinase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008652; P:amino acid biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001048; Aa_kinase.
DR InterPro; IPR002912; ACT.
DR InterPro; IPR001341; Aspartate_kinase.
DR InterPro; IPR005260; Asp_kin_monofn.
DR Pfam; PF00696; aak_kinase; 1.
DR Pfam; PF01842; ACT; 2.
DR TIGRFAMs; TIGR00657; asp_kinases; 1.
DR TIGRFAMs; TIGR00656; asp_kin_monofn; 1.
DR PROSITE; PS00324; ASPARTOKINASE; 1.
DR KINASE; Complete proteome.
SQ SEQUENCE 472 AA; 51022 MW; AFS130532F81A5F7 CRC64;

Query March 23.4%; Score 657.5; DB 17; Length 472;
Best Local Similarity 35.0%; Pred. No. 5.1e-42;
Matches 167; Conservative 95; Mismatches 182; Indels 33; Gaps 11;

QY 83 LSVYMKFGSSVSAPMAVAGLITPPE--RPVYVLSAMGKTTNNL-----LAGKA 136
DB 1 MKIVMKGGISVDGKKIRVAQLKRYHEGNIIVVTGALGVTDRLLENARLASTKG 60
QY 137 VGGGVHVSIEEM--NMVSLIKTVDELGLPXICNTSLY-----ELBQLKGIMM 187
DB 61 -----KSLVKEFKTELTKHKEHAKDAIEDPRAKEVLYQVLDRIEELKALIGICVL 114
QY 188 KEILPTSDYLVSPFGCMSTRIPSAVLNKRVAROYAPDIGFITTDSPGNADILENTY 247
DB 115 GEILTSRIDYISSYGERLAPIVSGAVRSIGASISIEYTGEGAGIVTSDYGNARPLEKTY 174
QY 248 PAVAKRLHGMIDDPAPVYTGFLGKMGSGAVTTIGRGSDLTATTIGKALGLREIOWM 307
DB 175 ELVYKRL--GRLIES--HLVVTGFIGEN--EDGITTTIGRSGSDPSASILGALKADEIMWM 231
QY 308 KDYDGVLTCDPNTYPAKIVPYLTFFBATELAYFGAQLHPQSMRPARRGDIPVRYKNSY 367
DB 232 KEVNGIWTDPRIAPFEAKTIPOISYAEAMELSYFGANVLHPRIIEPAKREHIFPVKNYF 291
QY 368 NPKAPGTLITRQDMDGIVLVLSIVKSNVTMLDIVSTRMLGOVGFARVSGICIEDL 427
DB 292 NPEPFGTLVVAEKQCHRVKAVSLI--KVALINISGABMPGTGTVALFTALARAQV 349
QY 428 CISYDCAVTSVSVSLDPSKIMSRELIOAGSELDHYVEELEKIAIVLLIOQRAIISLI 487
DB 350 NIWISQSSSESNLSFVVSSESHVSAALKALHAEFNREIYVEIISDRNV-----CVAVAV 403
QY 488 G-NVEGSSLLIEKGRILRKSGVNVQMSOGASKNMSLLVHSDAKALVEALHQAIF 543
DB 404 GAGAGTTPGVAKRVEGALGNSMTNITMISQSSQYINISFVREGDFAAFAVKTLDHDF 460
```